

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 18.0952 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-3  
Perfect score: 218  
Sequence: 1 SPLPGWEERQDILGRYYVNHESRRTQWKRPDPQNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pbp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pbp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pbp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pbp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pbp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pbp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	218	100.0	927	3	US-08-895-601-6
2	209	95.9	38	3	US-08-630-916A-18
3	203	93.1	38	2	US-09-066-074-11
4	203	93.1	38	2	US-08-555-912A-11
5	203	93.1	38	3	US-08-630-916A-17
6	203	93.1	38	3	US-08-348-518C-13
7	203	93.1	38	3	US-08-476-509B-13
8	203	93.1	38	4	US-09-252-404A-36
9	203	93.1	38	4	US-09-275-900-11
10	148	67.9	834	3	US-08-539-205A-6
11	148	67.9	834	4	US-09-392-163A-6
12	142	65.1	38	3	US-08-630-916A-36
13	138	63.3	38	3	US-08-630-916A-20
14	138	63.3	38	3	US-08-348-518C-14
15	138	63.3	38	3	US-08-476-509B-14
16	133	61.0	38	3	US-08-630-916A-21
17	132	60.6	38	2	US-09-066-074-12
18	132	60.6	38	2	US-08-555-912A-12
19	132	60.6	38	3	US-08-630-916A-19
20	132	60.6	38	3	US-08-348-518C-16
21	132	60.6	38	3	US-08-476-509B-16
22	132	60.6	38	4	US-09-252-404A-37
23	132	60.6	38	4	US-09-275-900-12
24	128.5	58.9	906	3	US-08-630-916A-48
25	128	58.7	38	3	US-08-630-916A-35
26	128	58.7	474	4	US-09-774-639-371
27	128	58.7	752	4	US-09-919-039-235

28	128	58.7	852	2	US-09-070-060-3	Sequence 3, Appli
29	128	58.7	852	3	US-09-357-746-3	Sequence 3, Appli
30	128	58.7	854	2	US-09-070-060-4	Sequence 4, Appli
31	128	58.7	854	3	US-09-357-746-4	Sequence 4, Appli
32	126	57.8	89	4	US-09-248-796A-18186	Sequence 18186, A
33	123	56.4	766	3	US-08-539-205A-4	Sequence 4, Appli
34	123	56.4	766	4	US-09-392-163A-4	Sequence 4, Appli
35	122	56.0	683	3	US-08-630-916A-46	Sequence 46, Appli
36	121	55.5	30	4	US-09-252-404A-3	Sequence 3, Appli
37	119	54.6	38	3	US-08-630-916A-25	Sequence 25, Appli
38	119	54.6	38	3	US-08-630-916A-32	Sequence 31, Appli
39	118	54.1	38	3	US-08-630-916A-24	Sequence 32, Appli
40	115	52.8	38	3	US-08-630-916A-15	Sequence 15, Appli
41	115	52.8	38	3	US-08-476-509B-15	Sequence 15, Appli
42	112	51.4	38	3	US-08-630-916A-30	Sequence 30, Appli
43	111	50.9	735	3	US-08-539-205A-2	Sequence 2, Appli
44	111	50.9	735	4	US-09-392-163A-2	Sequence 2, Appli
45	111	50.9	735	4	US-09-392-163A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-895-601-6  
; Sequence 6, Application US/08895501  
; Patent No. 6060262  
; GENERAL INFORMATION:  
; APPLICANT: Beer-Romero, Peggy  
; APPLICANT: Strack, Peter J.  
; APPLICANT: Glass, Susan J.  
; APPLICANT: Rolfe, Mark  
; TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,  
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/895,601  
; FILING DATE: 16-JUL-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MIV-096.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 927 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-895-601-6

Query Match 100.0%; Score 218; DB 3; Length 927;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SPLPGWEERQDILGRYYVNHESRRTQWKRPDPQNL 38  
DB 218 SPLPGWEERQDILGRYYVNHESRRTQWKRPDPQNL 255

```
RESULT 2
US-08-630-916A-18
; Sequence 18, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING MW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-916A-18

Query Match          95.9%; Score 209; DB 3; Length 38;
Best Local Similarity 97.4%; Pred. No. 6.2e-22;
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38
Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

RESULT 3
US-09-066-074-11
; Sequence 11, Application US/09066074
; Patent No. 5952467
; GENERAL INFORMATION:
; APPLICANT: Hunter et al., Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,074
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,074
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/555,912
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-066-074-11

Query Match          93.1%; Score 203; DB 2; Length 38;
Best Local Similarity 89.5%; Pred. No. 4.2e-21;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38
Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

RESULT 4
US-08-555-912A-11
; Sequence 11, Application US/08555912A
; Patent No. 5972697
; GENERAL INFORMATION:
; APPLICANT: Hunter et al., Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,912A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```



.....



Query Match 67.9%; Score 148; DB 4; Length 834;  
Best Local Similarity 71.4%; Pred. No. 6e-12;  
Matches 25; Conservative 3; Mismatches 7; Indels

STATE: New York

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
;

US-08-476-509B-14



THIS PAGE BLANK (user)

THIS PAGE BLANK (user)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 15.7143 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-2  
Perfect score: 129  
Sequence: 1 GPLPVGWEXXXXXXXHNTTXXWXPX 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	75.2	38	US-08-630-916A-36	Sequence 36, Appl
2	95	73.6	158	US-09-270-767-59345	Sequence 59345, A
3	95	73.6	435	US-09-270-767-43942	Sequence 43942, A
4	93	72.1	33	US-08-630-916A-39	Sequence 39, Appl
5	92	71.3	38	US-08-630-916A-26	Sequence 26, Appl
6	92	71.3	38	US-08-630-916A-35	Sequence 35, Appl
7	92	71.3	38	US-08-348-518C-18	Sequence 18, Appl
8	92	71.3	38	US-08-476-509B-18	Sequence 18, Appl
9	92	71.3	766	US-08-539-205A-4	Sequence 4, Appl
10	92	71.3	766	US-09-392-163A-4	Sequence 4, Appl
11	92	71.3	906	US-08-630-916A-48	Sequence 48, Appl
12	91	70.5	38	US-09-066-074-12	Sequence 12, Appl
13	91	70.5	38	US-08-555-912A-12	Sequence 12, Appl
14	91	70.5	38	US-08-630-916A-19	Sequence 19, Appl
15	91	70.5	38	US-08-630-916A-38	Sequence 38, Appl
16	91	70.5	38	US-08-348-518C-16	Sequence 16, Appl
17	91	70.5	38	US-08-476-509B-16	Sequence 16, Appl
18	91	70.5	38	US-09-252-404A-37	Sequence 37, Appl
19	91	70.5	38	US-09-275-900-12	Sequence 12, Appl
20	91	70.5	224	US-08-630-916A-50	Sequence 50, Appl
21	91	70.5	297	US-09-949-016-8946	Sequence 8946, Ap
22	90	69.8	38	US-08-630-916A-14	Sequence 14, Appl
23	90	69.8	38	US-08-630-916A-16	Sequence 16, Appl
24	90	69.8	38	US-08-348-518C-12	Sequence 12, Appl
25	90	69.8	38	US-08-476-509B-12	Sequence 12, Appl
26	90	69.8	472	US-08-348-518C-5	Sequence 5, Appl
27	90	69.8	472	US-08-476-509B-5	Sequence 5, Appl

28	90	69.8	486	3	US-08-348-518C-2	Sequence 2, Appl
29	89	69.0	38	3	US-08-630-916A-22	Sequence 22, Appl
30	89	69.0	38	3	US-08-348-518C-17	Sequence 17, Appl
31	89	69.0	38	3	US-08-476-509B-17	Sequence 17, Appl
32	89	69.0	474	4	US-09-774-639-371	Sequence 371, App
33	89	69.0	752	4	US-09-919-039-235	Sequence 235, App
34	89	69.0	852	2	US-09-070-060-3	Sequence 3, Appl
35	89	69.0	852	3	US-09-357-746-3	Sequence 3, Appl
36	89	69.0	854	2	US-09-070-060-4	Sequence 4, Appl
37	89	69.0	854	3	US-09-357-746-4	Sequence 4, Appl
38	88	68.2	38	3	US-08-630-916A-15	Sequence 15, Appl
39	88	68.2	38	3	US-08-630-916A-32	Sequence 32, Appl
40	88	68.2	38	3	US-08-348-518C-11	Sequence 11, Appl
41	88	68.2	38	3	US-08-476-509B-11	Sequence 11, Appl
42	88	68.2	683	3	US-08-630-916A-46	Sequence 46, Appl
43	87	67.4	89	4	US-09-248-796A-18186	Sequence 18186, A
44	87	67.4	834	3	US-08-539-205A-6	Sequence 6, Appl
45	87	67.4	834	4	US-09-392-163A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-630-916A-36  
; Sequence 36, Application US/08630916A  
; Patent No. 601137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-36

Query Match 75.2%; Score 97; DB 3; Length 38;  
Best Local Similarity 53.1%; Pred. No. 1.6e-10;  
Matches 17; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GPLPVGWEXXXXXXXHNTTXXWXP 32  
Db 1 GPLPVGWEXXXXXXXHNTTXXWXP 32



RESULT 5  
US-08-630-916A-26  
; Sequence 26, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/630.916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 26:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-26  
Query Match 71.3%; Score 92; DB 3; Length 38;  
Best Local Similarity 46.9%; Pred. No. 1.3e-09;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;  
Qy 1 GPLPQGWEXXXXXXGXYYXXHNTXTTXXXP 32  
||||| : ||||| :  
Db 1 GPLPSGWMRLTNTARVYFVDHNTKTTTWDPP 32  
RESULT 6  
US-08-630-916A-35  
; Sequence 35, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/630.916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-35  
Query Match 71.3%; Score 92; DB 3; Length 38;  
Best Local Similarity 51.6%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
Qy 2 PLPQGWEXXXXXXGXYYXXHNTXTTXXXP 32  
||||| : ||||| :  
Db 2 PLPPGWEKRTDPRGRFYVDHNTTWTQRP 32  
RESULT 7  
US-08-348-518C-18  
; Sequence 18, Application US/08348518C  
; Patent No. 6022740  
; GENERAL INFORMATION:  
; APPLICANT: SUDOL, MARIUS  
; APPLICANT: PEER, BORK  
; APPLICANT: HENRY, CHEN  
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN. A  
; SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
; PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/348.518C  
; FILING DATE: 01-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-3
; IMMEDIATE SOURCE:
; CLONE: Rsp5
US-08-348-518C-18

Query Match 71.3%; Score 92; DB 3; Length 38;
Best Local Similarity 46.9%; Pred. No. 1.3e-09;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 1 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 32

RESULT 9
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-4

Query Match 71.3%; Score 92; DB 3; Length 766;
Best Local Similarity 46.9%; Pred. No. 3.7e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 345 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 376

RESULT 10
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-3
; IMMEDIATE SOURCE:
; CLONE: Rsp5
US-08-348-518C-18

Query Match 71.3%; Score 92; DB 3; Length 38;
Best Local Similarity 46.9%; Pred. No. 1.3e-09;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 1 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 32

RESULT 9
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-4

Query Match 71.3%; Score 92; DB 3; Length 766;
Best Local Similarity 46.9%; Pred. No. 3.7e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 345 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 376

RESULT 10
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-3
; IMMEDIATE SOURCE:
; CLONE: Rsp5
US-08-348-518C-18

Query Match 71.3%; Score 92; DB 3; Length 38;
Best Local Similarity 46.9%; Pred. No. 1.3e-09;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 1 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 32

RESULT 9
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-4

Query Match 71.3%; Score 92; DB 3; Length 766;
Best Local Similarity 46.9%; Pred. No. 3.7e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEHXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
   : |||||
Db 345 GPLPSGWEMLTNTARVYFVDHNTKTTTWD DP 376

RESULT 10
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
```



```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-4

Query Match 71.3%; Score 92; DB 4; Length 766;
Best Local Similarity 46.9%; Pred. No. 3.7e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXYYXXHNTXTTXXXP 32
    ||||| : |||||
Db 345 GPLPSGWEMRLTNTARVYVDHNTKTTWDDP 376

RESULT 11
US-08-630-916A-48
; Sequence 48, Application US/08630916A
; Patent No. 601137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
```

```
;
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-916A-48

Query Match 71.3%; Score 92; DB 3; Length 906;
Best Local Similarity 51.6%; Pred. No. 4.4e-08;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPXGWEXXXXXXGXYYXXHNTXTTXXXP 32
    ||||| : |||||
Db 367 PLPPGWKRTDPRGRFYVDHNTTRTTWQRP 397

RESULT 12
US-09-066-074-12
; Sequence 12, Application US/09066074
; Patent No. 5952467
; GENERAL INFORMATION:
; APPLICANT: Hunter et al., Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,074
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/555,912
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: RSP5/SC
; US-09-066-074-12

Query Match 70.5%; Score 91; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 1.9e-09;
Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXYYXXHNTXTTXXXP 32
    ||||| : |||||
Db 1 GRPPGWERTDNGRTYYVDHNTTRTTWQRP 32

RESULT 13
US-08-555-912A-12
; Sequence 12, Application US/08555912A
; Patent No. 5972697
; GENERAL INFORMATION:
```

APPLICANT: Hunter et al., Tony  
TITLE OF INVENTION: NIMA INTERACTING PROTEINS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/555,912A  
FILING DATE: 13-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07251/011001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IMMEDIATE SOURCE:  
CLONE: RSP5/SC  
US-08-555-912A-12

Query Match 70.5%; Score 91; DB 2; Length 38;  
Best Local Similarity 50.0%; Pred. No. 1.9e-09;  
Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXHNTXTTXWXP 32  
Db 1 GRLLPGWERTDNFGRTYYVDHNTTTWKRP 32

RESULT 14  
US-08-630-916A-19  
Sequence 19, Application US/08630916A  
Patent No. 601137  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
APPLICANT: Kay, Brian K.  
APPLICANT: Fowlkes, Dana M.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,916A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 896-8864/9741  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-916A-19  
Query Match 70.5%; Score 91; DB 3; Length 38;  
Best Local Similarity 50.0%; Pred. No. 1.9e-09;  
Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXHNTXTTXWXP 32  
Db 1 GRLLPGWERTDNFGRTYYVDHNTTTWKRP 32

RESULT 15  
US-08-630-916A-38  
Sequence 38, Application US/08630916A  
Patent No. 601137  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
APPLICANT: Kay, Brian K.  
APPLICANT: Fowlkes, Dana M.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,916A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-916A-38  
Query Match 70.5%; Score 91; DB 3; Length 38;  
Best Local Similarity 46.9%; Pred. No. 1.9e-09;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXHNTXTTXWXP 32

Db           |||||   ||   |: ||| || | | |  
          1 GPLPENWEMAYTENGVEYFIDHNTKTSWLDP 32

Search completed: October 13, 2005, 14:03:04  
Job time : 15.8393 secs

THIS PAGE BLANK (0370)

THIS PAGE BLANK (0370)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 243.297 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-1

Perfect score: 271  
Sequence: 1 XXXXXXXXXXXXFWX.....SXXXXXXLXAXXXXXXXF 108

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140	51.7	358	7	ADD15147 Human 503
2	131	48.3	108	4	AAB83016 Human HEC
3	127	46.9	514	4	AAU87301 Novel cen
4	127	46.9	514	4	ADM20055 Protein e
5	127	46.9	514	8	ADI54616 Novel hum
6	127	46.9	748	2	AAW13384 Human pro
7	127	46.9	748	4	AAB31477 Amino aci
8	127	46.9	748	7	ADN95750 Human BEC
9	127	46.9	804	4	AAU19610 Human dia
10	127	46.9	831	8	ADN05464 Antipsori
11	125	46.1	156	6	ABU70611 Human adi
12	125	46.1	308	8	ADP54729 Human PRO
13	125	46.1	310	7	ADP62087 Rat prote
14	125	46.1	804	2	AAW13387 Fission y
15	125	46.1	2044	7	ADP31153 Human dia
16	125	46.1	2309	8	ADP11077 Human the
17	125	46.1	4374	7	ADP78997 Human pro
18	123.5	45.6	766	2	AAW13385 Human pro
19	122	45.0	1094	5	ABG70126 Human pre
20	122	45.0	1488	5	ABG70111 Human pre
21	122	45.0	2011	8	ADJ68961 Human hea
22	122	45.0	2011	8	ADQ44003 Amino aci
23	121	44.6	5002	4	ABE63723 Drosophil
24	120.5	44.5	870	4	AAE05495 Human ubi
25	120.5	44.5	870	6	AAE32722 Nedd-4-1i

ALIGNMENTS

RESULT 1

ADD15147

ID ADD15147 standard; protein; 358 AA.

AC ADD15147;

DT 15-JAN-2004 (first entry)

DE Human 50352 polypeptide E6-AP carboxyl terminus domain.

KW Human; ubiquitin-protein ligase 50352; neural tube defect;  
forebrain anomaly; posterior fossa anomaly; syringomyelia; colon tumour;  
cancer; cerebrovascular disease; hypoxia; ischaemia; infarction;  
cellular proliferative disorder; cellular differentiative disorder;  
hormonal disorder; immune disorder; inflammatory disorder;  
neurological disorder; blood vessel disorder; cytostatic; cardiac;  
vasotropic; endocrine; respiratory; antinflammatory; neuroprotective;  
enzyme; E6-AP carboxyl terminus domain.

OS Homo sapiens.

PN US2003100020-A1.

PD 29-MAY-2003.

PF 09-OCT-2002; 2002US-00268036.

PR 09-OCT-2001; 2001US-0327820P.

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers RE;

DR WPI; 2003-787017/74.

PT Novel ubiquitin-protein ligase family member 50352 useful as diagnostic  
targets and therapeutic agent for controlling 50352 associated disorders  
such as neural tube defects, perinatal brain injury, hypoxia, ischemia.

PS Disclosure; SEQ ID NO 5; 53pp; English.

XX The invention relates to the human ubiquitin-protein ligase 50352  
polypeptide and the polynucleotide encoding it. The polypeptide is useful  
for treating 50352-associated disorders such as neural tube defects,  
forebrain anomalies, posterior fossa anomalies, syringomyelia, colon  
tumours and cerebrovascular diseases such as hypoxia, ischaemia and  
infarction. The 50352 polypeptide and its modulators can act as  
therapeutic agents for controlling cellular proliferative and/or









PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241828P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249267P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
FA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX  
DR WPI; 2001-476159/51.  
DR N-PSDB; ADM19576.  
XX  
PT Isolated nucleic acid molecule encoding a channel/transporter protein is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
PS Claim 11; SEQ ID NO 862; 809pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule encoding a  
CC channel/transporter protein or sequences at least 95% identical to a  
CC these. The nucleic acids and proteins encoded by them are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. The antibodies to the proteins can also be used  
CC in alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.  
CC The polypeptides can also be used to aid wound healing and epithelial  
CC cell proliferation, to prevent skin aging due to sunburn, to maintain  
CC organs before transplantation, for supporting cell culture of primary  
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can  
CC also be used as a food additive or preservative to increase or decrease  
CC storage capabilities. This sequence corresponds to a protein of the  
CC invention.  
XX  
SQ Sequence 514 AA;  
Query Match 46.9%; Score 127; DB 4; Length 514;  
Best Local Similarity 32.9%; Pred. No. 7.8e-15;  
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;  
Qy - 17 WFWIXXXXXXEXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXLPPXH 76  
Dy 421 WFWKAVBFFDEBRARLLQFVTGSSRVLPQGFALQGAAGPRLFTIHOIDACTNNLPKAH 480  
Qy 77 TCFNXLDPYXXXXXXLXXAI 101

Db 481 TCFNRIDIPPYESYKLVKLLTAI 505

RESULT 5

ID ADI54616 standard; protein; 514 AA.

XX AC ADI54616;

DT 16-DEC-2004 (first entry)

XX DE Novel human protein seq id 819.

XX KW neuroprotective; neurotropic; antiparkinsonian; anticonvulsant;

KW antidiabetic; antirheumatic; antiarthritic; dermatological;

KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;

KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;

KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;

KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;

KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;

KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;

KW amyotrophic lateral sclerosis; multiple sclerosis;

KW immune system disorder; diabetes; rheumatoid arthritis;

KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;

KW inflammatory disorder; ischaemia-reperfusion injury;

KW inflammatory bowel disease; Crohn's disease; infectious disease;

KW HIV infection; hepatitis infection; bacterial infection;

KW fungal infection; parasitic infection; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;

KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;

KW renal disorder; acute glomerulonephritis; pyelonephritis;

KW renal lithiasis; proliferative disorder; cancerous diseases; human.

XX OS Homo sapiens.

XX US2004018969-A1.

XX 29-JAN-2004.

XX 17-JAN-2001; 2001US-00764875.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217497P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239335P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.



KW	Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;
KW	transforming growth factor-beta; human; TGF-beta; chondrogenesis;
KW	osteogenesis; blood differentiation; cartilage formation; hair growth;
KW	neural tube patterning; retinal development; heart induction;
KW	morphogenesis; tooth formation; gamete formation.
XX	
OS	Homo sapiens.
XX	
PN	WO200077168-A2.
XX	
PD	21-DEC-2000.
XX	
PF	12-JUN-2000; 200WO-US016250.
XX	
PR	11-JUN-1999; 99US-0138969P.
XX	
PA	(UNYNY ) UNIV NEW YORK STATE RES FOUND.
PA	(HSCR-) HSC RES & DEV LP.
XX	
PI	Thomsen GH, Wrana J;
XX	
DR	WPI; 2001-071267/08.
DR	N-PSDB; AAF24853.
XX	
PT	Novel isolated Smurf protein useful for inhibiting bone morphogenic
PT	protein or tumor growth factor-beta activation pathway, for treating
PT	cancer and to block osteogenesis, hair growth, tooth formation.
XX	
PS	Claim 10; Fig 12; 107pp; English.
XX	
CC	The present sequence represents a human Smurf2 polypeptide. The
CC	specification also describes a Smurf1 polypeptide. Smurf polypeptides are
CC	negative regulators of Smad signal transduction, and antagonists of bone
CC	morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)
CC	signalling pathway. Expression of Smurf1 in a cell is useful for
CC	inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
CC	polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
CC	differentiation, cartilage formation, neural tube patterning, retinal
CC	development, heart induction and morphogenesis, hair growth, tooth
CC	formation, gamete formation and a wide variety of tissue and organ
CC	formation processes, and hinder the regeneration, growth, maintenance,
CC	etc., of bone and other tissues that are dependent on the BMP pathway.
CC	The polypeptide is useful for screening for various drugs and/or
CC	antibodies that can either enhance the BMP pathway, or inhibit it
XX	
SQ	Sequence 748 AA;
Query Match 46.9%; Score 127; DB 4; Length 748;	
Best Local Similarity 32.9%; Pred. No. 1.1e-14;	
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;	
QY	17 WFWXIXXXXXXEXXXXXXQFTGXRLPXGXFXLXXXXXXXIXXXXXXXLPXHX 76
Db	655 WFWKAVFFDEERRARLLQFVTGSSRVLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAH 714
QY	77 TCFNKLDPXPYXKXXXXXXLXXAI 101
Db	715 TCFNRIDIPPYESYKLYEKLTLTAI 739
RESULT 8	
ID	ADN95750
ID	ADN95750 standard; protein; 748 AA.
AC	ADN95750;
XX	
DT	01-JUL-2004 (first entry)
XX	
DE	Human BEC/LEC-related protein sequence SeqID674.
XX	
KW	growth, differentiation; blood endothelial cell; BEC;
KW	lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW	lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cyostatic;

04-DEC-2001 (first entry)  
Human diagnostic and therapeutic polypeptide (DITHP) #196.  
Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
respiratory disorder.  
Homo sapiens.  
WO200162927-A2.  
30-AUG-2001.  
21-FEB-2001; 2001WO-US006059.  
24-FEB-2000; 2000US-0184693P.  
24-FEB-2000; 2000US-0184697P.  
24-FEB-2000; 2000US-0184698P.  
24-FEB-2000; 2000US-0184768P.  
24-FEB-2000; 2000US-0184769P.  
24-FEB-2000; 2000US-0184770P.  
24-FEB-2000; 2000US-0184771P.  
24-FEB-2000; 2000US-0184772P.  
24-FEB-2000; 2000US-0184773P.  
24-FEB-2000; 2000US-0184774P.  
24-FEB-2000; 2000US-0184776P.  
24-FEB-2000; 2000US-0184777P.  
24-FEB-2000; 2000US-0184779P.  
24-FEB-2000; 2000US-0184813P.  
24-FEB-2000; 2000US-0184837P.  
24-FEB-2000; 2000US-0184841P.  
24-FEB-2000; 2000US-0185213P.  
24-FEB-2000; 2000US-0185216P.  
12-MAY-2000; 2000US-0203785P.  
15-MAY-2000; 2000US-0204228P.  
16-MAY-2000; 2000US-0204525P.  
16-MAY-2000; 2000US-0204821P.  
16-MAY-2000; 2000US-0204908P.  
16-MAY-2000; 2000US-0205232P.  
17-MAY-2000; 2000US-0204815P.  
17-MAY-2000; 2000US-0204863P.  
17-MAY-2000; 2000US-0205221P.  
17-MAY-2000; 2000US-0205285P.  
17-MAY-2000; 2000US-0205286P.  
17-MAY-2000; 2000US-0205287P.  
17-MAY-2000; 2000US-0205287P.  
17-MAY-2000; 2000US-0205323P.  
17-MAY-2000; 2000US-0205324P.  
(INCY-) INCYTE GENOMICS INC.  
Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
Chen A, D'sa SA, Amshey S, Dahl CR, Dam TC, Daniels SE, Dufour GE;  
Flores V, Fong WT, Greenwalt JB, Hillman JL, Jones AL, Liu TP;  
Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;  
Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
WPI; 2001-502867/55.  
N-PSDB; AAS31181.  
Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
Claim 27; Page 512-514; 522pp; English.  
The invention relates to polynucleotides (I) encoding diagnostic and  
therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and  
proteins involved in growth and development and receptors. (I) and (II)  
may be used in the prevention, diagnosis and treatment of diseases  
associated with inappropriate DITHP expression. For example, (I) and (II)  
may be used to treat disorders associated with decreased polypeptide

expression by rectifying mutations or deletions in a patient's genome,  
that affect the activity of the DITHPs, by expressing inactive proteins  
or supplementing the patient's own production of them. (I) and (II) may  
be used to treat diseases, for example, cell proliferative disorder,  
Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
(I) may be used to produce the DITHPs, by inserting the nucleic acids  
into a host cell and culturing the cell to express the protein. (I) and  
its complementary sequences may also be used as DNA probes in diagnostic  
assays to detect and quantitate the presence of similar nucleic acids in  
samples, and therefore which patients may be in need of restorative  
therapy. (II) may also be used as antigens in the production of  
antibodies against DITHPs and in assays to identify modulators of DITHP  
expression and activity. The anti-DITHP antibodies and antagonists may  
also be used to down regulate expression and activity. The anti-DITHP  
antibodies may also be used as diagnostic agents for detecting the  
presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay  
(ELISA)). AAU19415-AAU19625 represent human diagnostic and therapeutic  
(DITHP) polypeptides of the invention  
SQ Sequence 804 AA;  
Query Match 46.9%; Score 127; DB 4; Length 804;  
Best Local Similarity 32.9%; Pred. No. 1.2e-14;  
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;  
Qy 17 WFWXIXXXXXXQFTGXRLPXXGFXLXXXXXXXIXXXXXXXLPXXH 76  
Db 711 WFWKAVEFDEERRARLLQFVTGSSRVPLQGFALQGAAGPRLFTIHQIDACTNNLPKHA 770  
Qy 77 TCFNXLDPYXXSXXXXXXLXXAI 101  
Db 771 TCFNRIDIPPYESYKLYEKLITAI 795  
RESULT 10  
ADN05464  
ID ADN05464 standard; protein; 831 AA.  
XX  
AC ADN05464;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Antipsoriatic protein sequence #900.  
XX  
KW antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX  
OS Homo sapiens.  
XX  
FN WO2004028479-A2.  
XX  
PD 08-APR-2004.  
XX  
PP 25-SEP-2003; 2003WO-US030907.  
XX  
PR 25-SEP-2002; 2002US-0414006P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;  
XX  
DR WPI; 2004-305105/28.  
DR N-PSDB; ADN05463.  
XX  
PT New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX  
PS Claim 9; SEQ ID NO 1858; 3069pp; English.  
XX  
CC The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the

The invention relates to a complex between two interacting proteins in adipocyte cells, given in the specification. The proteins are identified by selecting a bait protein from a known adipocyte marker and then performing a yeast 2-hybrid selection to isolate prey proteins encoded by members of an adipocyte cDNA library. The proteins are designated SID (RTM) (selected interacting domains) proteins. Also included are a polynucleotide encoding a polypeptide in the adipocyte cells, a recombinant host cell expressing at least one of the interacting polypeptides of the complex, selecting a modulating compound in adipocyte cells, a SID (RTM) polypeptide comprising any of the 738 amino acid sequences given in the specification (including its fragment or variant), a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences given in the specification (including its fragment or variant), a vector comprising the SID (RTM) polynucleotide, a recombinant host cell comprising the vector, a protein chip comprising the polypeptides and a record comprising all or part of the data, listed in the specification.

CC antagonist of the polypeptide or an antibody that binds to the  
 CC polypeptide in combination with a carrier; (8) an article of manufacture  
 CC comprising a container, a label on the container and a composition of  
 CC matter of (7); (9) a method of treating an immune related disease in a  
 CC mammal; (10) a method for determining the presence of a PRO polypeptide  
 CC in a sample suspected of having the polypeptide; (11) a method of  
 CC diagnosing an immune related disease or an inflammatory immune response  
 CC in a mammal; (12) a method of identifying a compound that inhibits or  
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
 CC ; and (13) a method of stimulating the immune response in a mammal. The  
 CC PRO sequences have anti-allergic, anti-inflammatory, antiarthritic,  
 CC antiasthmatic, antidiabetic, antianemic, antipneumonia, antipneumonia,  
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,  
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
 CC virucide activities, and can be used in gene therapy. The nucleic acid  
 CC (I) and the encoded polypeptides, compositions, kits and methods are  
 CC useful in diagnosing and treating an immune related disease and in  
 CC stimulating an immune response. The present sequence represents a human  
 CC PRO protein from the present invention.  
 CC  
 CC Sequence 308 AA;

Query Match 46.1%; Score 125; DB 8; Length 308;  
 Best Local Similarity 32.9%; Pred. No. 1.1e-14;  
 Matches 28; Conservative 2; Mismatches 55; Indels 0; Gaps 0;  
 Qy 17 WFWIXXXXXXQFTGXRLPXPXGXFXLXXXXXXXXXXXXXLPXHX 76  
 Db 214 WFWRALRSFDQADRAKFLQFVTGTSKVPLOGFAALGNGMIQKQFIHRDDRSTDLPSAH 273  
 Qy 77 TCFNKLDPXYXKXXXXXXLXXAI 101  
 Db 274 TCFNQLDPAYESFEKLRHMLLAI 298

RESULT 13  
 ADE62087  
 ID ADE62087 standard; protein; 310 AA.  
 AC ADE62087;  
 XX  
 XX 29-JAN-2004 (first entry)  
 DT  
 XX  
 XX Rat Protein AAA81950, SEQ ID NO 8016.  
 DE  
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 XX Rattus norvegicus.  
 OS  
 XX WO2003016475-A2.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX 14-AUG-2002; 2002WO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR  
 XX 01-NOV-2001; 2001US-0346382P.  
 PR  
 XX 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 PA  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; AAA81950.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 310 AA;

Query Match 46.1%; Score 125; DB 7; Length 310;  
 Best Local Similarity 32.9%; Pred. No. 1.1e-14;  
 Matches 28; Conservative 2; Mismatches 55; Indels 0; Gaps 0;  
 Qy 17 WFWIXXXXXXQFTGXRLPXPXGXFXLXXXXXXXXXXXXXLPXHX 76  
 Db 214 WFWRALRSFDQADRAKFLQFVTGTSKVPLOGFAALGNGMIQKQFIHRDDRSTDLPSAH 273  
 Qy 77 TCFNKLDPXYXKXXXXXXLXXAI 101  
 Db 274 TCFNQLDPAYESFEKLRHMLLAI 298

RESULT 14  
 AAW13387  
 ID AAW13387 standard; protein; 804 AA.  
 XX  
 AC AAW13387;  
 XX  
 XX 10-JUL-1997 (first entry)  
 DT  
 XX  
 XX Fission yeast protein ubiquitin ligase publ.  
 DE  
 XX Protein ubiquitin ligase; publi; cdc25 phosphatase; CDK kinase; p53;  
 KW cell cycle; transgenic animal.  
 XX  
 XX Schizosaccharomyces pombe.  
 OS  
 XX WO9712962-A1.  
 PN  
 XX 10-APR-1997.  
 PD  
 XX 04-OCT-1996; 96WO-US015930.  
 PF  
 XX 04-OCT-1995; 95US-00539205.  
 PR  
 XX (COLD-) COLD SPRING HARBOR LAB.  
 XX  
 XX Beach D, Caligiuri M, Nefsky B;  
 XX WPI; 1997-226206/20.  
 DR N-PSDB; AAT62067.





Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	127	46.9	735	3	US-08-539-205A-2	Sequence 2, Appli
2	127	46.9	735	4	US-09-392-163A-2	Sequence 2, Appli
3	123.5	45.6	766	3	US-08-539-205A-4	Sequence 4, Appli
4	123.5	45.6	766	4	US-09-392-163A-4	Sequence 4, Appli
5	121	44.6	119	4	US-09-270-767-31689	Sequence 31689, A
6	121	44.6	119	4	US-09-270-767-46906	Sequence 46906, A
7	120.5	44.5	906	3	US-08-530-916A-18	Sequence 48, Appli
8	119	43.9	384	4	US-09-270-767-45284	Sequence 45284, A
9	116.5	43.0	464	4	US-09-248-796A-18187	Sequence 18187, A
10	116.5	43.0	474	4	US-09-774-639-371	Sequence 371, App
11	116.5	43.0	752	4	US-09-919-039-235	Sequence 235, App
12	116.5	43.0	852	2	US-09-070-060-3	Sequence 3, Appli
13	116.5	43.0	852	3	US-09-357-746-3	Sequence 3, Appli
14	116.5	43.0	854	2	US-09-070-060-4	Sequence 4, Appli
15	116.5	43.0	854	3	US-09-357-746-4	Sequence 4, Appli
16	116.5	43.0	927	3	US-08-895-601-6	Sequence 6, Appli
17	113.5	41.9	834	3	US-08-539-205A-6	Sequence 6, Appli
18	113.5	41.9	834	4	US-09-392-163A-6	Sequence 6, Appli
19	110.5	40.8	102	4	US-09-270-767-45062	Sequence 45062, A
20	100.5	37.1	868	4	US-09-949-016-11723	Sequence 11723, A
21	100.5	37.1	874	3	US-08-247-904B-8	Sequence 8, Appli
22	100.5	37.1	874	3	US-08-767-942A-21	Sequence 21, Appli
23	99.5	36.7	866	1	US-08-100-692-1	Sequence 1, Appli
24	99.5	36.7	866	2	US-08-574-030-1	Sequence 1, Appli
25	93.5	34.5	91	4	US-09-270-767-33525	Sequence 33525, A
26	93.5	34.5	91	4	US-09-270-767-48742	Sequence 48742, A
27	82.5	30.4	68	4	US-09-270-767-32280	Sequence 32280, A

```
Db      702 TCFNRIDIPPYESYKLYEKLTLTAI 726

RESULT 2
US-09-392-163A-2
; Sequence 2, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-2

Query Match 46.9%; Score 127; DB 4; Length 735;
Best Local Similarity 32.9%; Pred. No. 4.7e-16;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy      17 WFWXIXXXXXXQFTGXRLPXXGFXLXXXXXXXIXXXXXXXLPXXH 76
Db      642 WFWKAVEFFDEERRARLLQVGTSSRVPLQGFKALQGAAGFRLLTHQIDACTNNLPKHAH 701

Qy      77 TCFNXLDPXXSXXXXXXLXXAI 101
Db      702 TCFNRIDIPPYESYKLYEKLTLTAI 726

RESULT 3
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston

Db      702 TCFNRIDIPPYESYKLYEKLTLTAI 726

Query Match 45.6%; Score 123.5; DB 3; Length 766;
Best Local Similarity 34.1%; Pred. No. 2.3e-15;
Matches 29; Conservative 3; Mismatches 52; Indels 1; Gaps 1;

Qy      17 WFWXIXXXXXXQFTGXRLPXXGFXLXXXXXXXIXXXXXXXLPXXH 76
Db      674 WFWELMDEWSNEKSKRLLOFTTGTSTRIPVNGFKDLQSGDGRKFTI-EKAGEPNKLPKHAH 732

Qy      77 TCFNXLDPXXSXXXXXXLXXAI 101
Db      733 TCFNRDLPPYTSKKDLDHKLSTAV 757

RESULT 4
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-539-205A-4
```

```
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-4

Query Match 45.6%; Score 123.5; DB 4; Length 766;
Best Local Similarity 34.1%; Pred. No. 2.3e-15;
Matches 29; Conservative 3; Mismatches 55; Indels 1; Gaps 1;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXXXLPPXH 76
Db 674 WFWELMDEWSNEKSRLLQFTTGTIRIPVNGFKDQSGDGRKFTI-EKAGEPNKLPAH 732

Qy 77 TCFNKLDPYKXXXXXXLXXAI 101
Db 733 TCFNRLDLPPTYKKDLHKLIAV 757

RESULT 5
US-09-270-767-31689
; Sequence 31689, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 31689
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31689

Query Match 44.6%; Score 121; DB 4; Length 119;
Best Local Similarity 32.9%; Pred. No. 9.2e-16;
Matches 28; Conservative 2; Mismatches 55; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXXXLPPXH 76
Db 25 WFWRALRSFDQADRAKFLQFTGTISKVPLQFGSLEGMNGIQKFIHRDDRSTDLPCA 84

Qy 77 TCFNKLDPYKXXXXXXLXXAI 101
Db 85 TCFNQLDLPYKSYDKLRSLCKAI 109

RESULT 6
US-09-270-767-46906
; Sequence 46906, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 46906
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46906

Query Match 44.6%; Score 121; DB 4; Length 119;
Best Local Similarity 32.9%; Pred. No. 9.2e-16;
Matches 28; Conservative 2; Mismatches 55; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXXXLPPXH 76
Db 25 WFWRALRSFDQADRAKFLQFTGTISKVPLQFGSLEGMNGIQKFIHRDDRSTDLPCA 84

Qy 77 TCFNKLDPYKXXXXXXLXXAI 101
Db 85 TCFNQLDLPYKSYDKLRSLCKAI 109

RESULT 7
US-08-630-916A-48
; Sequence 48, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-916A-48

Query Match 44.5%; Score 120.5; DB 3; Length 906;
Best Local Similarity 36.5%; Pred. No. 1.1e-14;
Matches 31; Conservative 1; Mismatches 52; Indels 1; Gaps 1;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXXXLPPXH 76
Db 814 WFOVVKEMDNEKIRLLQFTGTCTRLPVGGFABLSGNGPKFCI-DKVGKETWLPKSH 872

Qy 77 TCFNKLDPYKXXXXXXLXXAI 101
Db 873 TCFNRLDLPYKSYEQRLREKLLYAI 897

RESULT 8
US-09-270-767-45284
; Sequence 45284, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
```



Matches 31; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
Qy 17 WFWIXXXXXXQFTGXRLPXPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 660 WFWQFVKEIDNEKRMRLQFVTGTCRLPVGGFADLMGNGPQKFCI-EKVCKENWLP RSH 718  
Qy 77 TCFNXLDPXPYKXXXXXXLXXAI 101  
Db 719 TCFNRLDLPYKSYEQLEKLLFAI 743  
RESULT 12  
US-09-070-060-3  
; Sequence 3, Application US/09070060  
; Patent No. 5976849  
; GENERAL INFORMATION:  
; APPLICANT: Hustad, Carolyn M.  
; APPLICANT: Childval, Namit  
; TITLE OF INVENTION: Human E3 Ubiquitin Protein  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; STREET: ZENECA Pharmaceuticals, Inc.  
; STREET: 1800 Concord Pike  
; CITY: Wilmington  
; STATE: DE  
; COUNTRY: USA  
; ZIP: 19850-5437  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070.060  
; FILING DATE: 30-APR-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/073.839  
; FILING DATE: 05-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Higgins, Patrick H  
; REGISTRATION NUMBER: 39,709  
; REFERENCE/DOCKET NUMBER: PHM.70312  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302.886.4889  
; TELEFAX: 302.886.8221  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 852 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-09-070-060-3  
Query Match 43.0%; Score 116.5; DB 2; Length 852;  
Best Local Similarity 36.5%; Pred. No. 6e-14;  
Matches 31; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
Qy 17 WFWIXXXXXXQFTGXRLPXPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 760 WFWQFVKEIDNEKRMRLQFVTGTCRLPVGGFADLMGNGPQKFCI-EKVCKENWLP RSH 818  
Qy 77 TCFNXLDPXPYKXXXXXXLXXAI 101  
Db 819 TCFNRLDLPYKSYEQLEKLLFAI 843  
RESULT 13  
US-09-357-746-3  
; Sequence 3, Application US/09357746

; Patent No. 6087122  
; GENERAL INFORMATION:  
; APPLICANT: ZENECA Limited  
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE  
; FILE REFERENCE: PHM.70312.N1  
; CURRENT APPLICATION NUMBER: US/09/357.746  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073.839  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: US No. 608712209/070.060  
; EARLIER FILING DATE: 1998-04-30  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 852  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-357-746-3  
Query Match 43.0%; Score 116.5; DB 3; Length 852;  
Best Local Similarity 36.5%; Pred. No. 6e-14;  
Matches 31; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
Qy 17 WFWIXXXXXXQFTGXRLPXPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 760 WFWQFVKEIDNEKRMRLQFVTGTCRLPVGGFADLMGNGPQKFCI-EKVCKENWLP RSH 818  
Qy 77 TCFNXLDPXPYKXXXXXXLXXAI 101  
Db 819 TCFNRLDLPYKSYEQLEKLLFAI 843  
RESULT 14  
US-09-070-060-4  
; Sequence 4, Application US/09070060  
; Patent No. 5976849  
; GENERAL INFORMATION:  
; APPLICANT: Hustad, Carolyn M.  
; APPLICANT: Childval, Namit  
; TITLE OF INVENTION: Human E3 Ubiquitin Protein  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.  
; STREET: 1800 Concord Pike  
; CITY: Wilmington  
; STATE: DE  
; COUNTRY: USA  
; ZIP: 19850-5437  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070.060  
; FILING DATE: 30-APR-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/073.839  
; FILING DATE: 05-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Higgins, Patrick H  
; REGISTRATION NUMBER: 39,709  
; REFERENCE/DOCKET NUMBER: PHM.70312  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302.886.4889  
; TELEFAX: 302.886.8221  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 854 amino acids  
; TYPE: amino acid

```

; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-070-060-4

Query Match      43.0%; Score 116.5; DB 2; Length 854;
Best Local Similarity 36.5%; Pred. No. 6e-14;
Matches 31; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 17 WFWXXXXXXXXXXQFXTGXXRLPXXGFXLXXXXXXXXXXXXXXXXLPPXH 76
Db 762 WFWQFVKEIDNEKRMRLQFVTGTCRLPVGFGADLMGSGNGPQKFCI-EKVGKENWLPESH 820

QY 77 TCFNXLDPXXYXXXXXXLXXAI 101
Db 821 TCFNRLDLPYKSYEQLEKLLFAI 845

```

```

RESULT 15
US-09-357-746-4
; Sequence 4, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-357-746-4

```

```

Query Match      43.0%; Score 116.5; DB 3; Length 854;
Best Local Similarity 36.5%; Pred. No. 6e-14;
Matches 31; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 17 WFWXXXXXXXXXXQFXTGXXRLPXXGFXLXXXXXXXXXXXXXXXXLPPXH 76
Db 762 WFWQFVKEIDNEKRMRLQFVTGTCRLPVGFGADLMGSGNGPQKFCI-EKVGKENWLPESH 820

QY 77 TCFNXLDPXXYXXXXXXLXXAI 101
Db 821 TCFNRLDLPYKSYEQLEKLLFAI 845

```

Search completed: October 13, 2005, 14:03:04  
Job time : 52.5536 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 230.242 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-1

Perfect score: 271

Sequence: 1 XXXXXXXXXXXXXXXFWX.....XXXXXXXXXXAIXXXXXXF 108

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	51.7	358	14	US-10-268-036-5
2	140	51.7	358	16	US-10-391-364-33
3	138	50.9	301	16	US-10-425-115-250383
4	137	50.6	1111	16	US-10-425-115-250386
5	136	50.2	661	16	US-10-425-115-292299
6	136	50.2	936	15	US-10-424-599-196163
7	135	49.8	1104	15	US-10-424-599-196163
8	135	49.8	3647	16	US-10-437-963-119793
9	132	48.7	1843	15	US-10-424-599-146004
10	131	48.3	108	14	US-10-307-956-1
11	127	46.9	514	11	US-09-764-875-819

12	127	46.9	735	14	US-10-313-955-2	Sequence 2, Appli
13	127	46.9	748	14	US-10-021-660-81	Sequence 81, Appl
14	127	46.9	748	18	US-10-756-149-4725	Sequence 4725, Ap
15	127	46.9	804	15	US-10-320-120-407	Sequence 407, App
16	123	45.6	766	14	US-10-313-955-4	Sequence 4, Appli
17	122	45.0	1094	14	US-10-043-487-300	Sequence 300, App
18	122	45.0	1488	14	US-10-043-487-285	Sequence 285, App
19	122	45.0	2011	16	US-10-408-765A-767	Sequence 767, App
20	121	44.6	5002	20	US-11-097-143-17961	Sequence 17961, A
21	120.5	44.5	870	14	US-10-097-534-12	Sequence 12, Appl
22	120.5	44.5	870	16	US-10-723-860-2167	Sequence 2167, Ap
23	120.5	44.5	906	14	US-10-185-050-48	Sequence 48, Appl
24	120	44.3	823	14	US-10-097-534-25	Sequence 25, Appl
25	120	44.3	823	18	US-10-756-149-5703	Sequence 5703, Ap
26	119.5	44.1	722	14	US-10-097-534-14	Sequence 14, Appl
27	119	43.9	1035	20	US-11-097-143-10152	Sequence 10152, A
28	118.5	43.7	898	15	US-10-188-186-114	Sequence 114, App
29	118.5	43.7	930	16	US-10-618-408-2	Sequence 2, Appli
30	117.5	43.4	949	20	US-11-097-143-5685	Sequence 5685, Ap
31	117.5	43.4	949	20	US-11-097-143-6360	Sequence 6360, Ap
32	116.5	43.0	474	10	US-09-774-639-371	Sequence 371, App
33	116.5	43.0	474	10	US-09-969-730-249	Sequence 249, App
34	116.5	43.0	474	15	US-10-621-363-249	Sequence 249, App
35	116.5	43.0	739	14	US-10-097-534-10	Sequence 10, Appl
36	116.5	43.0	739	15	US-10-374-979-89	Sequence 89, Appl
37	116.5	43.0	739	15	US-10-182-936A-89	Sequence 89, Appl
38	116.5	43.0	739	16	US-10-477-238A-668	Sequence 668, App
39	116.5	43.0	739	16	US-10-680-287A-668	Sequence 668, App
40	116.5	43.0	739	17	US-10-477-173-668	Sequence 668, App
41	116.5	43.0	739	18	US-10-450-763-46836	Sequence 46836, A
42	116.5	43.0	752	10	US-09-919-039-235	Sequence 235, App
43	116.5	43.0	832	14	US-10-032-585-7296	Sequence 7296, Ap
44	116.5	43.0	854	14	US-10-287-218-3	Sequence 3, Appli
45	116.5	43.0	854	16	US-10-474-291-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-10-268-036-5

; Sequence 5, Application US/10268036

; Publication No. US20030100020A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Meyers, Rachel E.

; TITLE OF INVENTION: 50352, A HUMAN UBIQUITIN-PROTEIN LIGASE

; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR

; FILE REFERENCE: MPI01-224PIRM

; CURRENT APPLICATION NUMBER: US/10/268,036

; CURRENT FILING DATE: 2003-10-09

; PRIOR APPLICATION NUMBER: US 60/327,820

; PRIOR FILING DATE: 2001-10-09

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Consensus

US-10-268-036-5

Query Match	51.7%	Score 140;	DB 14;	Length 358;
Best Local Similarity	37.6%	Pred. No. 1.9e-17;		
Matches	32;	Conservative	0;	Mismatches 53;
				Indels 0;
				Gaps 0;
Qy	17	WFMXXXXXXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXLPPXXH	76	
Db	264	WFMEIVEEFDEERAKLQFVTGSPRLPQGFKSLESGNGIPKFTIEKAGTERLPTAH	323	
Qy	77	TCFNKLDLPYXXSXXXXXXXXXXAI	101	

Db 324 TCFNRDLDPKYSKSEILRSKLLAI 348

RESULT 2

US-10-391-364-33

Sequence 33, Application US/10391364

Publication No. US20040121349A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.

APPLICANT: Meyers, Rachel E.

APPLICANT: Carroll, Joseph M.

APPLICANT: Cook, William James

APPLICANT: Kapeller-Libermann, Rosana

APPLICANT: Weich, Nadine S.

APPLICANT: Bandaru, Rajasekhar

TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,

TITLE OF INVENTION: 16558, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND

TITLE OF INVENTION: US957 THEREFOR

FILE REFERENCE: MPI03-0190WNIM

CURRENT APPLICATION NUMBER: US/10/391,364

CURRENT FILING DATE: 2003-03-18

PRIOR APPLICATION NUMBER: US 09/950,370

PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: US 60/231,084

PRIOR FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: US 10/294,039

PRIOR FILING DATE: 2002-11-13

PRIOR APPLICATION NUMBER: US 60/338,587

PRIOR FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 10/266,035

PRIOR FILING DATE: 2002-10-07

PRIOR APPLICATION NUMBER: US 60/328,198

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: US 09/717,926

PRIOR FILING DATE: 2000-11-21

PRIOR APPLICATION NUMBER: US 60/214,707

PRIOR FILING DATE: 2000-06-27

PRIOR APPLICATION NUMBER: US 10/268,036

PRIOR FILING DATE: 2002-10-09

PRIOR APPLICATION NUMBER: US 60/327,820

PRIOR FILING DATE: 2001-10-09

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 93

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 33

LENGTH: 358

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Consensus sequence

US-10-391-364-33

Query Match 51.7%; Score 140; DB 16; Length 358;

Best Local Similarity 37.6%; Pred. No. 1.9e-17;

Matches 32; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 17 WFWXIXXXXXXXXQFXTGXRLPXXGFXLXXXXXXXXXXXXXLPXXH 76

Db 264 WFWVEEFPEKBRKLLQFVTGSPRLPQGFSGNGIPKFTIEKAGTEDERLPTAH 323

Qy 77 TCFNRDLDPKYSKSEILRSKLLAI 101

Db 324 TCFNRDLDPKYSKSEILRSKLLAI 348

RESULT 3

US-10-425-115-250383

Sequence 250383, Application US/10425115

Publication No. US2004021427A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 33

LENGTH: 1111

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Consensus sequence

US-10-425-115-250386

Query Match 50.6%; Score 137; DB 16; Length 1111;

Best Local Similarity 33.7%; Pred. No. 2.1e-16;

Matches 31; Conservative 2; Mismatches 59; Indels 0; Gaps 0;

Qy 17 WFWXIXXXXXXXXQFXTGXRLPXXGFXLXXXXXXXXXXXXXLPXXH 76

Db 1017 WFWVEIVQGFSGKEDKARFLQFVTGTSKVPLEGFSELGISGQRFQIHKAYGSTNHLPSAH 1076

Qy 77 TCFNRDLDPKYSKSEILRSKLLAI 108

Db 1077 TCFNRDLDPKYSKSEILRSKLLAI 1108

RESULT 4

US-10-425-115-250386

Sequence 250386, Application US/10425115

Publication No. US2004021427A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 250386

LENGTH: 1111

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Consensus sequence

US-10-425-115-250386

Query Match 50.9%; Score 138; DB 16; Length 301;

Best Local Similarity 33.7%; Pred. No. 3.9e-17;

Matches 31; Conservative 2; Mismatches 59; Indels 0; Gaps 0;

Qy 17 WFWXIXXXXXXXXQFXTGXRLPXXGFXLXXXXXXXXXXXXXLPXXH 76

Db 207 WFWVEIVQGFSGKEDKARFLQFVTGTSKVPLEGFSELGISGQRFQIHKAYGSTNHLPSAH 266

Qy 77 TCFNRDLDPKYSKSEILRSKLLAI 108

Db 267 TCFNRDLDPKYSKSEILRSKLLAI 298

RESULT 5

US-10-425-115-292299

Sequence 292299, Application US/10425115

Publication No. US2004021427A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 250386

LENGTH: 1111

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Consensus sequence

US-10-425-115-250386

Query Match 50.6%; Score 137; DB 16; Length 1111;

Best Local Similarity 33.7%; Pred. No. 2.1e-16;

Matches 31; Conservative 2; Mismatches 59; Indels 0; Gaps 0;

Qy 17 WFWXIXXXXXXXXQFXTGXRLPXXGFXLXXXXXXXXXXXXXLPXXH 76

Db 1017 WFWVEIVQGFSGKEDKARFLQFVTGTSKVPLEGFSELGISGQRFQIHKAYGSTNHLPSAH 1076

Qy 77 TCFNRDLDPKYSKSEILRSKLLAI 108

Db 1077 TCFNRDLDPKYSKSEILRSKLLAI 1108



```
Query Match      50.2%; Score 136; DB 15; Length 936;  
Best Local Similarity   32.6%; Pred.No. 2.7e-16;  
Matches          30; Conservative    3; Mismatches     59; Indels       0; Gaps        0;
```

-----

Qy	17	WFVAIXXXXXXXXQQFTGXRRLPDXGFXLXXXXXXXXXXXXXXXXXXLPXH 76     :     ::   :
Dd	842	WFEVQGSLSKEDKARLLQLFVTGTSTKVPLEGSALQTSGSQKFIIHKAYGPSDHPSAH 901     :     ::
Qy	77	TCFNXLDPYXKXXXXXXXXXXLAIXXXXXXP 108     :     ::
Dd	902	TCFNOLDPYPSPSKOHLLEERLLLAIHEASEGF 933     :     ::

```

RESULT 8
US-10-437-963-119793
; Sequence 119793, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 119793
; LENGTH: 3647
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_22976C.1 pep
US-10-437-963-119793

```

Query Match	49.8%	Score 135;	DB 16;	Length 3647;
Best Local Similarity	32.6%	Pred. No. 1.5e-15;		
Matches 30; Conservative	3;	Mismatches 59;	Indels 0;	Gaps 0;



```

; LOCATION: (81)...(81)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (86)...(86)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (88)...(88)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (90)...(92)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (93)...(93)
; OTHER INFORMATION: Xaa = Leucine or Methionine
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (94)...(95)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (96)...(96)
; OTHER INFORMATION: Xaa = Arginine or Lysine
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (98)...(99)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (102)...(107)
; OTHER INFORMATION: Xaa = any amino acid; 0-2 residues may be missing
US-10-307-956-1

Query Match 48.3%; Score 131; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXXLPXXH 76
Db 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXXLPXXH 76
Qy 77 TCFNXLDPYXYSXXXXXXLXXAI 101
Db 77 TCFNXLDPYXYSXXXXXXLXXAI 101

RESULT 11
US-09-764-875-819
; Sequence 819, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 819
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-819

Query Match 46.9%; Score 127; DB 11; Length 514;
Best Local Similarity 32.9%; Pred. No. 7.3e-15;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXXLPXXH 76

US-10-313-955-2
; Sequence 2, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Caligiuri, Maureen
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313,955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-313-955-2

Query Match 46.9%; Score 127; DB 14; Length 735;
Best Local Similarity 32.9%; Pred. No. 1e-14;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXXLPXXH 76
Db 642 WFWKAVEFFDEERRARLLQFVTGSSRVLQGFKALQGAAGPRLFTIHQIDACTNNLPKAAH 701
Qy 77 TCFNXLDPYXYSXXXXXXLXXAI 101
Db 702 TCFNRIDIPPYESYEKLYEKLTLTAI 726

RESULT 13
US-10-021-660-81
; Sequence 81, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
```

```

; APPLICANT: Glynnne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152326a1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021.660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-81

Query Match          46.9%; Score 127; DB 14; Length 748;
Best Local Similarity 32.9%; Pred. No. 1e-14;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWKIXXXXXXXXQFTGXRLPPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 655 WFWKAVEFFDEERRARLLQFVTGSSRVPLOGFKALQGAAGPRLFTIHQIDACTNNLPK 714

Qy 77 TCFNXLDPYXXSXXXXXXXXXXAI 101
Db 715 TCFNRIDIPPYESYKLYKLLTAI 739

RESULT 14
US-10-756-149-4725
; Sequence 4725, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756.149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4725
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4725

Query Match          46.9%; Score 127; DB 18; Length 748;
Best Local Similarity 32.9%; Pred. No. 1e-14;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWKIXXXXXXXXQFTGXRLPPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 655 WFWKAVEFFDEERRARLLQFVTGSSRVPLOGFKALQGAAGPRLFTIHQIDACTNNLPK 714

Qy 77 TCFNXLDPYXXSXXXXXXXXXXAI 101
Db 715 TCFNRIDIPPYESYKLYKLLTAI 739

RESULT 15
US-10-220-120-407
; Sequence 407, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKDREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220.120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,813; 60/184,773; 60/184,698; 60/184,770; 60/184,774;
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
; 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 407
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:132147.3.orf3:2000FEB18
US-10-220-120-407

Query Match          46.9%; Score 127; DB 15; Length 804;
Best Local Similarity 32.9%; Pred. No. 1.1e-14;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWKIXXXXXXXXQFTGXRLPPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 711 WFWKAVEFFDEERRARLLQFVTGSSRVPLOGFKALQGAAGPRLFTIHQIDACTNNLPK 770
Qy 77 TCFNXLDPYXXSXXXXXXXXXXAI 101
```

Db 771 TCFNRIDIPPYESYEKLYEKLITAI 795

Search completed: October 13, 2005, 14:49:58  
Job time : 232.242 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**







```
Matches 28; Conservative 2; Mismatches 55; Indels 0; Gaps 0;
QY 17 WFWIXIXXXXXXQFTGXRLRPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 214 WFWRALRSFDQADRAKFLQFTGTGSKVPLQGFALLEGNGIQQKFIHRDRSDRDLPSAH 273
QY 77 TCFNXLDPYXXSXXXXXXXXLXXAI 101
Db 274 TCFNQLDLPAYESFEKLRHMLLAI 298

RESULT 7
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37900
R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <R1>
A:Cross-references: UNIPROT:Q9UTG2; EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:
A:Experimental source: strain 972h-; cosmid c1805
C:Genetics:
A:Gene: SPDB:SPAC1805.15c
A:Map position: 1
A:Introns: 60/2, 105/1; 639/2
F:242-279/Domain: WW repeat homology <WWR>

Query Match 45.6%; Score 123.5; DB 2; Length 671;
Best Local Similarity 34.1%; Pred. No. 2.8e-15;
Matches 29; Conservative 3; Mismatches 55; Indels 1; Gaps 1;
QY 17 WFWIXIXXXXXXQFTGXRLRPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 579 WFWELLSWSPEKKAKLLQFATGTSRLPLSGFKDMHSGDGRPKFTI-EKVGHISQLPKAH 637
QY 77 TCFNXLDPYXXSXXXXXXXXLXXAI 101
Db 638 TCFNRLDLPYNSKEELQKLTIAI 662

RESULT 8
S66562
ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: E6-AP-like protein ubiquitin ligase
C:Species: Schizosaccharomyces pombe
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S66562; T45159
R:Nefsky, B.; Beach, D.
EMBO J. 15, 1301-1312, 1996
A:Title: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
A:Reference number: S66562; MUID:96205868; PMID:8635463
A:Accession: S66562
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-766 <NEP>
A:Cross-references: UNIPROT:Q92462; GB:U66716; NID:GL519443; PIDN:AAB07514.1; PID:G15194
R:Nefsky, B.S.; Beach, D.
submitted to the EMBL Data Library, August 1996
A:Description: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of
A:Reference number: Z22935
A:Accession: T45159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-766 <NE2>
A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1
C:Genetics:
A:Gene: pubi
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
```

```
C:Keywords: ligase
F:205-242/Domain: WW repeat homology <WW1>
F:288-325/Domain: WW repeat homology <WW2>
F:345-382/Domain: WW repeat homology <WW3>

Query Match 45.6%; Score 123.5; DB 1; Length 766;
Best Local Similarity 34.1%; Pred. No. 3.1e-15;
Matches 29; Conservative 3; Mismatches 52; Indels 1; Gaps 1;
QY 17 WFWIXIXXXXXXQFTGXRLRPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 674 WFWELMDWSNEKKSRLLOFTTGTSRIPVNGFKDLOGSDGRPKFTI-EKAGEPNKLPKAH 732
QY 77 TCFNXLDPYXXSXXXXXXXXLXXAI 101
Db 733 TCFNRLDLPYTSKKDLHKLIAV 757

RESULT 9
T37545
ubiquitin-protein ligase (EC 6.3.2.19) pubi [validated] - fission yeast (Schizosaccharom
N:Alternate names: ubiquitin ligase Pubi
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37545; T48655
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21722
A:Accession: T37545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <MCL>
A:Cross-references: UNIPROT:Q92462; EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:
A:Experimental source: strain 972h-; cosmid c11G7
R:Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.
Mol. Gen. Genet. 254, 520-538, 1997
A:Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pubi ubi
A:Reference number: Z07985; MUID:97340937; PMID:9197411
A:Accession: T48655
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <SAL>
A:Cross-references: EMBL:U62795; NID:g2262192; PIDN:AAB63350.1; PID:g2262193
A:Experimental source: strain J227
C:Genetics:
A:Gene: pubi; SPDB:SPAC11G7.02
A:Map position: 1
A:Introns: 6/2; 14/1; 62/2
C:Function: <UBI>
A:Description: EC 6.3.2.19 [validated, MUID:96205868]
C:Function: <CYC>
A:Description: involved in of the mitotic activating tyrosine phosphatase cdc25 [validat
C:Function: <TOL>
A:Description: required for low pH-tolerance [validated, MUID:97340937]
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C:Keywords: cell cycle control; ligase
F:205-242/Domain: WW repeat homology <WW1>
F:288-325/Domain: WW repeat homology <WW2>
F:345-382/Domain: WW repeat homology <WW3>

Query Match 45.6%; Score 123.5; DB 2; Length 767;
Best Local Similarity 34.1%; Pred. No. 3.1e-15;
Matches 29; Conservative 3; Mismatches 52; Indels 1; Gaps 1;
QY 17 WFWIXIXXXXXXQFTGXRLRPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 675 WFWELMDWSNEKKSRLLOFTTGTSRIPVNGFKDLOGSDGRPKFTI-EKAGEPNKLPKAH 733
QY 77 TCFNXLDPYXXSXXXXXXXXLXXAI 101
Db 734 TCFNRLDLPYTSKKDLHKLIAV 758
```



F:342-379/Domain: WW repeat homology <WWR2>  
F:393-430/Domain: WW repeat homology <WWR3>  
F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 41.1%; Score 111.5; DB 2; Length 820;  
Best Local Similarity 32.9%; Pred. No. 2.7e-13;  
Matches 28; Conservative 3; Mismatches 53; Indels 1; Gaps 1;

Qy 17 WFWIXXXXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 723 FFWQTVRSWDEQKSRLLQFTTGTSRIPVNGFKLQSGDGPFRFTI-EKAGEITNLPKAH 781

Qy 77 TCFNXLDPXYXXSXXXXXXXXXXAI 101  
Db 782 TCFNRLDLPYPYKSLMLYKLLTAV 806

Search completed: October 13, 2005, 15:11:12  
Job time : 47.0989 secs

F:342-379/Domain: WW repeat homology <WWR2>  
F:393-430/Domain: WW repeat homology <WWR3>  
F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 41.9%; Score 113.5; DB 2; Length 820;  
Best Local Similarity 32.9%; Pred. No. 2.7e-13;  
Matches 28; Conservative 3; Mismatches 53; Indels 1; Gaps 1;

Qy 17 WFWIXXXXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 727 WFWKAVLLMDAEKRIRLLQFTVGTSRVPMNGFAELYGSGNGQLFTI-PQWGSPEKLPAH 785

Qy 77 TCFNXLDPXYXXSXXXXXXXXXXAI 101  
Db 786 TCFNRLDLPYPYTFEFLREKLIMAV 810

RESULT 14  
T51886  
hypothetical protein DKFp434N1131.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C:Accession: T51886  
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, July 2000  
A:Reference number: Z25854  
A:Accession: T51886  
A:Molecule type: mRNA  
A:Status: preliminary  
A:Residues: 1-221 <AAA>  
A:Cross-references: UNIPROT:Q9NFS9; EMBL:AL390186  
A:Experimental source: adult testis; clone DKFp434N1131  
C:Genetics:  
A:Note: DKFp434N1131.1

Query Match 41.5%; Score 112.5; DB 2; Length 221;  
Best Local Similarity 31.8%; Pred. No. 1.1e-13;  
Matches 27; Conservative 3; Mismatches 54; Indels 1; Gaps 1;

Qy 17 WFWIXXXXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76  
Db 129 WFWAVERFNEQRLRLQFTVGTSSIPYGFASLRSNGSPRRFCV-EKWGKITALPAH 187

Qy 77 TCFNXLDPXYXXSXXXXXXXXXXAI 101  
Db 188 TCFNRLDLPYPYPSFMSLYEKLITAV 212

RESULT 15  
T49744  
probable ubiquitin-protein ligase [imported] - Neurospora crassa  
N:Alternate names: protein B24B19.160  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C:Accession: T49744  
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49744  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-815 <SCH>  
A:Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.160  
A:Experimental source: BAC clone B24B19; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B24B19.160  
A:Map position: 6  
A:Introns: 11/1; 24/1; 59/2; 110/1; 783/2  
C:Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology  
F:239-276/Domain: WW repeat homology <WWR1>  
F:334-371/Domain: WW repeat homology <WWR2>  
F:393-430/Domain: WW repeat homology <WWR3>

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 205.714 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-1  
Perfect score: 271  
Sequence: 1 XXXXXXXXXXXXXXXFWX.....SXXXXXXXXXXAIXXXXXX 108

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135	49.8	3684	1 UPL1_ARATH	Q8y23 arabidopsis
2	135	49.8	3716	2 Q6K3Y7	Q6K3Y7 oryza sativ
3	134	49.4	3658	1 UPL2_ARATH	Q8h0t4 arabidopsis
4	132	48.7	3349	2 Q6BJ34	Q6bj34 debaromyce
5	131	48.3	4177	2 Q9GUP2	Q9gup2 caenorhabdi
6	130	48.0	4065	1 TOM1_NEUCR	Q9p4z1 neurospora
7	129	47.6	3227	1 PRL1_SCHPO	O13834 schizosacch
8	129	47.6	3320	2 Q6CFL1	Q6cfl1 yarrowia li
9	128	47.2	295	2 Q8NDD8	Q9cun6 mus musculu
10	128	47.2	619	1 SUF1_MOUSE	Q8k300 mus musculu
11	128	47.2	728	2 Q8K300	Q8k300 mus musculu
12	127	46.9	133	2 Q80WX3	Q80wx3 mus musculu
13	127	46.9	190	2 Q6P066	Q6p066 mus musculu
14	127	46.9	258	2 Q9CSE3	Q9cse3 mus musculu
15	127	46.9	288	2 Q6GDE7	Q9ghau4 homo sapien
16	127	46.9	748	1 SUF2_HUMAN	Q9hau4 homo sapien
17	127	46.9	3242	2 Q6CWS8	Q6cws8 kluyveromyc
18	126.5	46.7	786	1 PUB3_SCHPO	O14326 schizosacch
19	125	46.1	159	2 Q9CU76	Q9cu76 mus musculu
20	125	46.1	321	2 Q6AXC1	Q6axc1 mus musculu
21	125	46.1	322	1 URB1_RAT	P51593 rattus norv
22	125	46.1	2749	1 URB1_MOUSE	Q7tmy8 mus musculu
23	125	46.1	3360	1 URB1_HUMAN	Q7z6z7 homo sapien
24	124	45.8	3258	1 TOM1_ASHGO	Q75692 ashbya gos
25	123.5	45.6	671	1 PUB2_SCHPO	Q9ut92 schizosacch
26	123.5	45.6	767	1 PUB1_SCHPO	Q92462 schizosacch
27	123	45.4	983	2 Q7PSQ3	Q7psq3 anopheles g
28	121	44.6	964	2 Q8T3L0	Q8t3l0 drosophila
29	121	44.6	3247	2 Q6FMP7	Q6fmp7 candida gla
30	121	44.6	5146	2 Q9VXR3	Q9vxr3 drosophila
31	120.5	44.5	108	2 Q9CT73	Q9ct73 mus musculu

32	120.5	44.5	870	1 WWP2_HUMAN	O0308 homo sapien
33	120.5	44.5	870	1 WWP2_MOUSE	Q8dbh0 mus musculu
34	120	44.3	597	2 Q80YC4	Q80yc4 mus musculu
35	120	44.3	823	2 Q7LDY1	Q7ldy1 homo sapien
36	120	44.3	823	2 Q6P9Q1	Q6p9q1 mus musculu
37	120	44.3	823	2 Q8C5W5	Q8c5w5 mus musculu
38	120	44.3	826	2 Q150J3	O150j3 homo sapien
39	120	44.3	826	2 Q8CHG5	Q8chg5 mus musculu
40	119.5	44.1	355	2 Q8BSC0	Q8bsc0 m mus muscu
41	119.5	44.1	731	1 SUF1_XENLA	Q8pun2 xenopus lae
42	119.5	44.1	757	1 SUF1_HUMAN	Q9hce7 homo sapien
43	119	43.9	1061	2 Q3V8E3	Q9v8e3 drosophila
44	119	43.9	2410	2 Q8SR14	Q8sr14 encephalico
45	118.5	43.7	252	2 Q7TMG3	Q7tmg3 mus musculu

ALIGNMENTS

RESULT 1  
UPL1\_ARATH  
ID UPL1\_ARATH STANDARD; PRT; 3684 AA.  
AC Q8GV23; Q9LG27; Q9M7K7;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE E3 ubiquitin protein ligase UPL1 (EC 6.3.2.-) (Ubiquitin-protein  
ligase 1).  
GN Name=UPL1; OrderedLocusNames=Atlg55860; ORFNames=F14J16.14, F14J16.37;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A., PROBABLE FUNCTION, TISSUE SPECIFICITY,  
RP DEVELOPMENTAL STAGE, AND MUTAGENESIS OF CYS-3648.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20040063; PubMed=10571878;  
RA Bates P.W., Vierstra R.D.;  
RT "UPL1 and 2, two 405 kDa ubiquitin-protein ligases from Arabidopsis  
thaliana related to the HECT-domain protein family";  
RL Plant J. 20:183-195 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
Largin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana";  
RL Nature 408:816-820 (2000).  
RN [3]  
RP SEQUENCE OF 3533-3681 FROM N.A.  
RC STRAIN=cv. Columbia;  
RX PubMed=11910074; DOI=10.1126/science.1071006;  
RA Seki M., Narusaka M., Kamiya A., Ishida J., Satou M., Sakurai T.,  
Nakashima M., Enju A., Akiyama K., Oono Y., Muramatsu M.,  
Hayashizaki Y., Kawai J., Carninci P., Itoh Y., Arakawa T.,





```
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchon S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boissrame A., Boyer J., Catolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382139; CAG90246.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR010314; DUF908.
DR InterPro; IPR010314; DUF913.
DR InterPro; IPR000569; HECT.
DR Pfam; PF06012; DUF908; 1.
DR Pfam; PF06025; DUF913; 1.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
SQ SEQUENCE 3349 AA; 363078 MW; A42D8924F2637D78 CRC64;

Query Match 48.7%; Score 132; DB 2; Length 3349;
Best Local Similarity 30.4%; Pred. No. 1.4e-15;
Matches 28; Conservative 4; Mismatches 60; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXLPXXH 76
Db 3255 WFWRAVXSFNDEERAKULLQFATGTSKVLPLNGFKELSGASGTCKFSIHRDYGTDLPLSSH 3314
Qy 77 TCFNKLDPYXSYXXXXXXLXAXIXXXXXX 108
Db 3315 TCFNQIDLPAVESYETLRGSLVLLAITEGHEGF 3346

RESULT 5
ID Q9GUP2 PRELIMINARY; PRT; 4177 AA.
AC Q9GUP2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y67D8C.5.
GN Name=Y67D8C.5; ORFNames=Y67D8C.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Edwards J., Du H., Lamar B., Kemp K., Wohldmann P., Walker C.;
RT "The sequence of C. elegans cosmid Y67D8C.";
```

```
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025724; AAC23375.2; -.
DR HSSP; Q9H0M0; IND7.
DR WormBase; WBGene0002069; Y67D8C.5.
DR WormPep; Y67D8C.5; CE31665.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR010314; DUF913.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR000449; UBA.
DR Pfam; PF06012; DUF908; 1.
DR Pfam; PF06025; DUF913; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS50030; UBA; 1.
KW Hypothetical protein.
SQ SEQUENCE 4177 AA; 465406 MW; 9C32EF90AB87FD58 CRC64;

Query Match 48.3%; Score 131; DB 2; Length 4177;
Best Local Similarity 34.1%; Pred. No. 2.7e-15;
Matches 29; Conservative 2; Mismatches 54; Indels 0; Gaps 0;
```

Qy 17 WFWIXXXXXXEXXXXXXQFTGXRLPXPXGFXLXXXXXXXIXXXXXXXLPXXH 76



```
Db 4083 WFWALRSFEKDKAKFLQFTGTSTKVPLOQFASLEGNGVQKPSIHWDNRGGRLPAAH 4142
Qy 77 TCFNKLDPXYXKXXXXXXLXXAI 101
Db 4143 TCFNQLDLPQYESYKLRQSLLLAI 4167

RESULT 6
TOM1_NEUCR
ID TOM1_NEUCR STANDARD; PRT; 4065 AA.
AC O9P4Zi;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE E3 ubiquitin protein ligase TOM1-like protein (EC 6.3.2.-).
GN ORFNames=B11B22.010, NCU08501.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Algn V.,
RA Hoheisel J.D., Fartmann B., Nyakatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence.";
RL Nucleic Acids Res. 31:1944-1954 (2003).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels P., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Tanakiev P., Ball-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysysselis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Kryzstofova S.,
RA Rasmussen C., Metzner R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Eboile D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 422:859-868 (2003).
CC -!- FUNCTION: Probable ubiquitin ligase protein, which may be involved
CC in mRNA export. E3 ubiquitin ligase protein mediate ubiquitination
CC and subsequent proteasomal degradation of target proteins.
CC Participates in mRNA export from the nucleus by regulating the
CC transport of hnRNP proteins (By similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the TOM1/PTR1 family.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL356834; CAB92704.2; -;
CC EMBL; AABX01000134; EAA34194.1; -;
CC
CC
```

```
DR PTR; T49799; T49799.
DR HSP; Q05086; IC42.
DR InterPro; IPR010309; DUF908.
DR InterPro; IPR010314; DUF913.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF06012; DUF908; 1.
DR Pfam; PF06025; DUF913; 1.
DR Pfam; PF06632; HECT; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS0237; HECT; 1.
KW Coiled coil; Ligase; mRNA transport; Nuclear protein; Transport;
KW Ubl conjugation pathway.
FT DOMAIN 3729 4065 HECT.
FT DOMAIN 1263 1287 Coiled coil (Potential).
FT DOMAIN 2491 2568 Coiled coil (Potential).
FT DOMAIN 2839 2924 Coiled coil (Potential).
FT DOMAIN 3320 3369 Coiled coil (Potential).
FT BINDING 4032 4032 Ubiquitin (By similarity).
SQ SEQUENCE 4065 AA; 452568 MW; F74683CEC36F9350 CRC64;

Query Match 48.0%; Score 130; DB 1; Length 4065;
Best Local Similarity 31.5%; Pred. No. 4.1e-15;
Matches 29; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

Qy 17 WFWXXXXXXXXXXXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
Db 3971 WFWRAVSFDKEALAKLQFVTGTSTKVPLOQFASLEGNGVSRFNIHRDYGSKDRLPSSH 4030
Qy 77 TCFNKLDPXYXKXXXXXXLXXAIXXXXXXF 108
Db 4031 TCFNQLDLPYENYETLRSQLLKAITAGSDYF 4062

RESULT 7
PTR1_SCHPO
ID PTR1_SCHPO STANDARD; PRT; 3227 AA.
AC O13834;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE E3 ubiquitin protein ligase ptr1 (EC 6.3.2.-) (Poly(A) + RNA transport
DE protein 1).
GN Name=ptr1; ORFNames=SPAC19D5.04;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Schaefer M., Mueller-Auer S.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs W., Fritzc C., Holzer E., Moesl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
```

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of *Schizosaccharomyces pombe*.";  
RL Nature 415:871-880 (2002).  
RN [2]  
RP FUNCTION, AND MUTANT PTR1-1.  
RX PubMed=15094387; DOI=10.1016/j.bbrc.2004.03.171;  
RA Andoh T., Azad A.K., Shigematsu A., Ohshima Y., Tani T.;  
RT "The fission yeast ptr1+ gene involved in nuclear mRNA export encodes  
a putative ubiquitin ligase.";  
RL Biochem. Biophys. Res. Commun. 317:1138-1143 (2004).  
CC -1- FUNCTION: Probable ubiquitin ligase protein involved in mRNA  
export. E3 ubiquitin ligase protein mediate ubiquitination and  
CC subsequent proteasomal degradation of target proteins. Probably  
CC participates in mRNA export from the nucleus by regulating the  
CC transport of hnRNP proteins such as rael.  
CC -1- PATHWAY: Ubiquitin conjugation; third step.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: Belongs to the TOM1/PTR1 family.  
CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z99531; CAB16714.1; -;  
DR PIR; T37964; T37964.  
DR HSP; Q05086; IC4Z.  
DR GeneDB Spombe; SPAC19D5.04; -;  
DR InterPro; IPR010309; DUF908.  
DR InterPro; IPR010314; DUF913.  
DR InterPro; IPR000569; HECT\_domain.  
DR Pfam; PF06012; DUF908; 1.  
DR Pfam; PF06025; DUF913; 1.  
DR Pfam; PF06032; HECT; 1.  
DR SMART; SM00119; HECTC; 1.  
DR PROSITE; PS50237; HECT; 1.  
DR Ligate; mRNA transport; Nuclear protein; Transport;  
KW Ub1 conjugation pathway.  
FT DOMAIN 2891 3227 HECT.  
FT DOMAIN 1854 2017 Asp-rich.  
FT BINDING 3194 3194 Ubiquitin (By similarity).  
FT MUTAGEN 2887 L->Q: In ptr1-1; induces defects in mRNA  
FT export.  
SQ SEQUENCE 3227 AA; 365028 MW; 07FC47AB79124575 CRC64;  
Query Match 47.6%; Score 129; DB 1; Length 3227;  
Best Local Similarity 31.5%; Pred. No. 5.1e-15;  
Matches 29; Conservative 3; Mismatches 60; Indels 0; Gaps 0;  
QY 17 WFWXIXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXIXXXXXXLPXXH 76  
DB 3133 WFWRAVSFDEERAKLLQVATGTSKVPPLNGFLEGMGSGFRFNTHKSYGSLNRLPQSH 3192  
QY 77 TCFNXLDPYXSXXXXXLLXAIXXXXXXF 108  
DB 3193 TCFNQLDLPDYDYEQLRSMLLTAINEGSEGF 3224

DE ubiquitin ligase (Fragment).  
GN ORFNames=YALI0B059409;  
OS Yarrowia lipolytica CLIB99.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=284591;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genolevures;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
RA Sweeney D., Tekaiia F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT "Genome evolution in yeasts.";  
RL Nature 430:35-44 (2004).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RA Genoscope;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CR382128; CAG82782.1; -;  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR000225; Armadillo.  
DR InterPro; IPR010309; DUF908.  
DR InterPro; IPR010314; DUF913.  
DR InterPro; IPR000569; HECT.  
DR Pfam; PF06012; DUF908; 1.  
DR Pfam; PF06025; DUF913; 1.  
DR Pfam; PF06032; HECT; 1.  
DR SMART; SM00119; HECTC; 1.  
DR PROSITE; PS50176; ARM\_REPEAT; 1.  
DR PROSITE; PS50237; HECT; 1.  
KW Ligate.  
FT NON TER  
SQ SEQUENCE 3320 AA; 369711 MW; CFC8C54DFA8C8C20 CRC64;  
Query Match 47.6%; Score 129; DB 2; Length 3320;  
Best Local Similarity 32.8%; Pred. No. 5.2e-15;  
Matches 30; Conservative 2; Mismatches 60; Indels 0; Gaps 0;  
QY 17 WFWXIXXXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXXIXXXXXXLPXXH 76  
DB 3226 WFWRAVSFDEERAKLLQVATGTSKVPPLDGFLEGMGSGFRFNTHRAIGNRLPSSH 3285  
QY 77 TCFNXLDPYXSXXXXXLLXAIXXXXXXF 108  
DB 3286 TCFNQLDLPDYDSVETLRGSLLLAITEGREGF 3317

RESULT 9  
Q8NDD8  
ID Q8NDD8 PRELIMINARY; PRT; 295 AA.  
AC Q8NDD8;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein DKFZp564H223 (fragment).  
GN Name=DKFZp564H223;





```

RESULT 13
Q6P066
ID Q6P066 PRELIMINARY; PRT; 190 AA.
AC Q6P066;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065796; AAH65796.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTc; 1.
DR PROSITE; PS0237; HECT; 1.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 190 AA; 22024 MW; 9D9996BF033EE17B CRC64;
Query Match 46.9%; Score 127; DB 2; Length 190;
Best Local Similarity 32.9%; Pred. No. 7.6e-16;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;
QY 17 WFWIXXXXXXEXXXXXXXQFTGXRLPXGFXLXXXXXXXXXXXXXXXXXXLPXXH 76
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97 WFWKAVEFDEERRARLLQFTVTSRVPLOGFKALQGAAGPRFTIHOIDACTNNLPKHA 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 77 TCFNKLDPXXKXXXXXXLXXAI 101
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 TCFNRIDIPPYESYEKYLELTAI 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
Q9CSE3
ID Q9CSE3 PRELIMINARY; PRT; 258 AA.
AC Q9CSE3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2810411E22 product:similar to SIMILAR TO E3
DE UBIQUITIN LIGASE SMURF2 (Fragment).
CN Name=Smurf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body; DOI=10.1016/S0076-6879(99)03004-9;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Segabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013082; BAB28637.1; -.
DR HSSP; Q9H0M0; 1ND7.
DR MGD; MGI:1913563; Smurf2.

```

```
DR GO: 0005622; C:intracellular; IEA.
DR GO: 0016874; F:ligase activity; IEA.
DR GO: 0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: 0006512; F:ubiquitin cycle; IEA.
DR InterPro; IPR00569; HECT.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECT; 1.
DR PROSITE; PS50237; HECT; 1.
KW Ligase.
FT NON_TER 1 1
SQ SEQUENCE 258 AA; 29670 MW; B0BE04BCB1FCBC98 CRC64;

Query Match 46.9%; Score 127; DB 2; Length 258;
Best Local Similarity 32.9%; Pred. No. 1e-15;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXLPPXH 76
Db 165 WFWKAVEFFDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAH 224

Qy 77 TCFNXLDPXXYXXXXXXLXXAI 101
Db 225 TCFNRIDIPPYESYEKLYEKLTAI 249

RESULT 15
Q96DE7 PRELIMINARY; PRT; 288 AA.
AC Q96DE7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SMURF2 protein (Fragment).
GN Name=SMURF2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haehn F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC009527; AA09527.1; -.
DR HSP; O9HOM0; IND7.
DR GO: 0005622; C:intracellular; IEA.
DR GO: 0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: 0006512; F:ubiquitin cycle; IEA.
DR InterPro; IPR00569; HECT.
```

```
DR SMART; SM00119; HECT; 1.
DR PROSITE; PS50237; HECT; 1.
FT NON_TER 1 1
SQ SEQUENCE 288 AA; 33255 MW; FE2B43E300B66537 CRC64;

Query Match 46.9%; Score 127; DB 2; Length 288;
Best Local Similarity 32.9%; Pred. No. 1.e-15;
Matches 28; Conservative 3; Mismatches 54; Indels 0; Gaps 0;

Qy 17 WFWIXXXXXXQFTGXRLPXXGFXLXXXXXXXXXXXXLPPXH 76
Db 195 WFWKAVEFFDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAH 254

Qy 77 TCFNXLDPXXYXXXXXXLXXAI 101
Db 255 TCFNRIDIPPYESYEKLYEKLTAI 279

Search completed: October 13, 2005, 15:09:12
Job time : 208.714 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 74.3407 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-2

Perfect score: 129  
Sequence: 1 GPLPQGWGXXXXXXGXXYYXHXNTTXXWXXPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	76.0	33	4	AAB83017 Human HEC
2	97	75.2	37	2	AAW37638 Peptide c
3	97	75.2	38	5	AAU87962 Human WW
4	97	75.2	38	7	ADB49230 Novel hum
5	95	73.6	1082	4	ABB63355 Drosophil
6	93	72.1	33	2	AAW38107 WW domain
7	93	72.1	33	7	ADB49233 WW domain
8	93	72.1	759	6	ABJ25504 Aspergill
9	93	72.1	869	6	ABJ26104 Aspergill
10	92	71.3	38	2	AAW97685 Yeast Rsp
11	92	71.3	38	2	AAW37637 Peptide c
12	92	71.3	38	4	AAW83024 Yeast Rsp
13	92	71.3	38	5	AAU87968 Human WW
14	92	71.3	38	7	ADB49229 Novel hum
15	92	71.3	38	7	ADB49220 Yeast WW
16	92	71.3	766	2	AAW13385 Human pro
17	92	71.3	832	5	ABP73459 Candida a
18	92	71.3	870	4	AAE05495 Human ubi
19	92	71.3	870	6	AAE32722 Nedd-4-li
20	92	71.3	870	8	ADQ19348 Human sof
21	92	71.3	870	8	ABM81770 Tumour-as
22	92	71.3	906	2	AAW36795 Novel hum
23	92	71.3	906	7	AAW36795 Novel hum
24	91	70.5	38	2	AAW37683 Yeast Rsp
25	91	70.5	38	2	AAW37640 Peptide c

ALIGNMENTS

RESULT 1

AAAB83017  
ID AAB83017 standard; peptide; 33 AA.

XX AC AAB83017;

DT 25-JUN-2001 (first entry)

DE Human HECT E3 ubiquitin ligase WW domain consensus sequence.

XX Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; WW domain.  
XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 5 /note= "Xaa = any amino acid"

FT Misc-difference 9..11 /note= "Xaa = any amino acid"

FT Misc-difference 12..14 /label= Ser, His, Pro, Asp, Glu, Thr, Tyr

FT /note= "each residue is an independently selected polar amino acid"

FT Misc-difference 16 /label= Ser, His, Pro, Asp, Glu, Thr, Tyr

FT /note= "independently selected polar amino acid"

FT Misc-difference 17 /note= "Xaa = any amino acid"

FT Misc-difference 20 /label= Ile, Val, Leu, Met

FT /note= "hydrophobic residue"

FT Misc-difference 21 /note= "Xaa = any amino acid"

FT Misc-difference 25 /label= Ser, His, Pro, Asp, Glu, Thr, Tyr

FT /note= "independently selected polar amino acid"

FT Misc-difference 28 /label= Ser, His, Pro, Asp, Glu, Thr, Tyr

FT /note= "independently selected polar amino acid"

FT Misc-difference 30 /note= "Xaa = any amino acid"

FT Misc-difference 31 /label= Ser, His, Pro, Asp, Glu, Thr, Tyr

```

FT Misc-difference 33 /note= "independently selected polar amino acid"
FT /label= Ser, His, Pro, Asp, Glu, Thr, Tyr
FT /note= "independently selected polar amino acid"
XX WO200116604-A1.
XX 08-MAR-2001.
XX 29-AUG-2000; 2000WO-US023729.
XX 30-AUG-1999; 99US-00385918.
XX (SIGN-) SIGNAL PHARM INC.
XX Hoekstra MF, Xie W, Murray BW, Mercurio FM;
XX WPI; 2001-327913/34.
XX Screening for modulators of TGF-beta and/or bone morphogenic protein
XX (BMP) mediated signaling useful for treating cancer and osteoporosis by
XX evaluating the ability of agents to modulate Smad protein degradation.
XX Claim 2; Page 34; 75pp; English.
XX The present sequence is the HECT (homologous to E6 carboxyl terminus) E3
XX ubiquitin ligase WW domain. The WW domain binds to the Smad PY motif,
XX resulting in ubiquitination of Smad by the E3 ubiquitin ligase. The
XX sequence is provided in a specification relating to a method for
XX screening for agents that modulate transforming growth factor (TGF)-beta
XX and/or bone morphogenic protein (BMP)-mediated signalling. The method
XX involves evaluating the effect of an agent on binding of HECT E3
XX ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad
XX protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein
XX HECT E3 ubiquitin ligase activity. The method is useful for stimulating
XX bone formation in a patient or treating a condition associated with
XX insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that
XX inhibit BMP-mediated signalling are useful for treating inflammation,
XX ageing, cancer and infectious diseases. Agents that augment BMP-mediated
XX signalling are useful for stimulating bone anabolism as well as treating
XX broken bones, osteoporosis, and acute or chronic renal failure. Agents
XX that inhibit TGF-mediated signalling are useful for treating cancer,
XX inflammation, neurodegeneration and fibrosis
XX SQ Sequence 33 AA;
Query Match 76.0%; Score 98; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPLPFGWEXXXXXXXGXXYYXXHNTXTTXWXXP 32
Db 1 GPLPFGWEXXXXXXXGXXYYXXHNTXTTXWXXP 32
RESULT 2
AAW37638
ID AAW37638 standard; peptide; 37 AA.
XX AAW37638;
XX 23-APR-1998 (first entry)
XX Peptide containing a WW domain.
XX WW domain; cell signalling; growth regulation; cytoskeleton organisation;
XX targeted drug screening; modulator; WW domain interaction; WWP4.
XX Synthetic.
XX WO9737223-A1.
XX 09-OCT-1997.
XX
/
XX PF 03-APR-1997; 97WO-US005547.
XX 03-APR-1996; 96US-00630916.
XX (CYTO-) CYTOGEN CORP.
XX (UYNC-) UNIV NORTH CAROLINA.
XX Pirozzi G, Kay BK, Fowlkes DM;
XX WPI; 1997-503234/46.
XX Identifying cell signalling and growth regulatory polypeptides by
XX reaction with multivalent recognition complex - polypeptides are useful
XX in targeted drug selection.
XX Claim 49; Page 104; 220pp; English.
XX Peptides AAW36798-801 and AAW37636-40 are peptides that contain WW
XX domains. The WW domain is a small functional domain found in a large
XX number of proteins from a variety of species including humans, nematodes
XX and yeast. Its name is derived from the observation that two tryptophan
XX residues, one in the amino terminal portion of the WW domain and one in
XX the carboxyl terminal portion, are conserved. Most proteins containing WW
XX domains have a function involving cell signalling and growth regulation
XX or the organisation of the cytoskeleton. Polypeptides containing a WW
XX domain are identified by treating a multivalent recognition unit complex
XX that has selective binding affinity for a WW domain, with many
XX polypeptides and identifying those with selective affinity for the
XX complex. Proteins containing WW domains are used for targeted drug
XX screening, i.e. to identify potential modulators of specific WW domain
XX interactions. The valency of the recognition unit is important in
XX determining specificity of interaction with WW domains. In multivalent
XX form specificity is relaxed, but not lost, so proteins containing WW
XX domains similar, but not identical, to the sequence of the peptides,
XX target WW can be detected, including new polypeptides
XX SQ Sequence 37 AA;
Query Match 75.2%; Score 97; DB 2; Length 37;
Best Local Similarity 53.1%; Pred. No. 9e-10;
Matches 17; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 GPLPFGWEXXXXXXXGXXYYXXHNTXTTXWXXP 32
Db 1 GPLPFGWEXXXXXXXGXXYYXXHNTXTTXWXXP 32
RESULT 3
AAU87962
ID AAU87962 standard; peptide; 38 AA.
XX AAU87962;
XX 05-JUN-2002 (first entry)
XX Human WW domain #15.
XX Human; PDZ domain; WW domain; rat; cow; mouse; fruitfly; protein therapy;
XX gene therapy; PDZ-mediated disease; inward potassium channel; WBP;
XX dimer inhibitor peptide; carboxylate binding loop.
XX Homo sapiens.
XX WO200207751-A1.
XX 31-JAN-2002.
XX 24-JUL-2001; 2001WO-US023269.
XX 25-JUL-2000; 2000US-0221215P.
XX 28-NOV-2000; 2000US-00723810.
XX

```







CC activities, or as an immunogen to generate antibodies. This is the amino  
CC acid sequence of a WW domain binding protein associated protein. Note:  
CC This sequence appears in the sequence listing but is not further  
CC described in the specification.

Sequence 33 AA;

Query Match 72.1%; Score 93; DB 7; Length 33;  
Best Local Similarity 70.0%;  
Pred. NO. 4.1e-09;  
Matches 21; Conservative 0; Mismatches 9; Indels

[illegible]

RESULT 8  
ABJ25504  
ID ABJ25504 standard; protein: 759 AA.

AC ABJ25504;

DT 16-APR-2003 (first entry)

DE Aspergillus fumigatus essential gene protein #162.

AA Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
KW cancer; contamination; biofilm; antibody; immune response.  
KW

OS *Aspergillus fumigatus*.

PN WO200286090-A2.

31-OCT-2002.

23-APR-2002

23-APP-2001-2001US-0285697P

PR 27-APR-2001; 2001US-0287066P.  
PR 05-JUN-2001; 2001US-0295890P

PR 09-JUL-2001; 2001US-0303899P.  
PR 31-AUG-2001; 2001US-0316362P

XX PA (E.I.T.-) ELITRA PHARM INC

XX Jiang B. Tishkoff D. Za

XX  
DB WPT: 2003-093124/08

New purified or isolated

PT or for treating a non-infectious disease in a subject e.g. cancer.

PS Disclosure: Page: 175pp: English.

The invention relates to novel purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as *A. fumigatus*, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by *A. fumigatus*, or to prevent or inhibit formation on a surface of a biofilm comprising *A. fumigatus*. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of *A. fumigatus* to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein

with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or elicit immune response as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This sequence represents a protein of one of the essential genes of *Aspergillus fumigatus* of the invention.

Sequence 759 AA;

Query Match 72.1%; Score 93; DB 6; Length 759;  
Best Local Similarity 50.0%; Pred. No. 1.3e-07;  
Matches 16; Conservative 0; Mismatches 16; Indels

**Qy** 1 GLPLPGWEXXXXXXGXYYXXHNTXTTXXXP 32  
| | | | | | | | | | | | | |  
**Pb** 151 GRPAGWEREDNLGRTVYVDHNTRTTTWTRP 182

RESULT 9  
ABJ26104

ID ABJ26104 standard; protein; 869 AA.

AC ABJ26104:

DT 16-APR-2003 (first entry)

XX DE *Aspergillus fumigatus* essential gene protein #762

XX	Fungicide; cytostatic; essential gene; <i>Aspergillus fumigatus</i> ; infection;
KW	cancer; contamination; biofilm; antibody; immune response
KW	

XX *Aspergillus fumigatus*

XX  
PN  
W0200286090-A2XX  
31-OCT-2003

XX  
DE  
23-APR-2002.

XX  
PR 23-APR-2001: 2001US-0285697P

PR 27-APR-2001; 2001US-0287066P.  
PR 05-JUN-2001; 2001US-0295890P

PR 09-JUL-2001; 2001US-0303899P.  
PR 31-AUG-2001; 2001US-0316362P

XX  
PA (E.L.T.T.) E.L.T.T.A PHARM INC

XX  
PT Jiang B Tishkoff D Zamudio C Eroshkin AM Hu W Lemieux SM.

XX  
DR WPT: 2003-093124/08

XX New purified or isolated nucleic acids of essential genes of *Aspergillus*  
PT *fumigatus*, useful for treating or preventing infections by *A. fumigatus*,  
PT or for treating a non-infectious disease in a subject e.g. cancer.

XX  
PS Disclosure: page: 175pp: English.

The invention relates to novel purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as *A. fumigatus*, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by *A. fumigatus*, or to prevent or inhibit formation on a surface of a biofilm comprising *A. fumigatus*. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of *A. fumigatus* to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and

CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This sequence represents a protein of one of the essential genes  
CC of Aspergillus fumigatus of the invention  
XX  
SQ Sequence 869 AA;

Query Match 72.1%; Score 93; DB 6; Length 869;  
Best Local Similarity 50.0%; Pred. No. 1.5e-07;  
Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
QY 1 GPLPGXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 283 GRLPAGWERREDNLGRYYVDHNTRTTTWTRP 314

RESULT 10  
AAR97685  
ID AAR97685 standard; protein; 38 AA.  
XX  
AC AAR97685;  
XX  
DT 30-AUG-1996 (first entry)  
XX  
DE Yeast Rsp5 WW domain-3.  
XX  
KW WW domain; signal transduction; diagnosis; gene therapy;  
KW Yes proto-oncogene associated protein; YAP; Rsp5.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN WO9617061-A1.  
XX  
PD 06-JUN-1996.  
XX  
PF 30-NOV-1995; 95WO-US015512.  
XX  
PR 01-DEC-1994; 94US-00348518.  
PR 07-JUN-1995; 95US-00476509.  
XX  
PA (UYRQ ) UNIV ROCKEFELLER.  
PA (DELB-) DELBRUECK CENT MOLECULAR MEDICINE MAX.  
XX  
PI Sudol M, Chen H, Bork P;  
XX  
DR WPI; 1996-286829/29.  
XX  
PT DNA encoding Yes proto-oncogene associated protein - used to modulate  
PT intracellular signal transduction e.g. for treatment of muscular  
PT dystrophy.  
XX  
PS Claim 10; Fig 12; 126pp; English.  
XX  
CC WW domains (AAR97673-92) were identified in a number of proteins involved  
CC in signalling or regulatory functions. The WW domain was initially  
CC identified in the res proto-oncogene associated proteins (YAPs) of  
CC chicken, human and mouse (see also AAR97669-70 and AAR97672). A consensus  
CC sequence is given in AAR97671. Yeast Rsp5 contains an N-terminal  
CC regulatory domain common to protein kinases C and synaptotagmins. The Rsp5  
CC WW domains (AAR97683-85) can be expressed in bacterial, yeast, insect or  
CC mammalian cells, and used to identify WW domain ligands. They can be  
CC introduced into cells, either directly or by gene therapy, to increase  
CC the level of signal transduction  
XX  
SQ Sequence 38 AA;

Query Match 71.3%; Score 92; DB 2; Length 38;  
Best Local Similarity 46.9%; Pred. No. 7.2e-09;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;  
  
QY 1 GPLPGXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 GPLPGXGWEMRLTNTARVYFVDHNTKTTTWDPP 32  
| | | | | | | | | | | | | | | | | | | | | |  
  
RESULT 11  
AAW37637  
ID AAW37637 standard; peptide; 38 AA.  
XX  
AC AAW37637;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Peptide containing a WW domain.  
XX  
KW WW domain; cell signalling; growth regulation; cytoskeleton organisation;  
KW targeted drug screening; modulator; WW domain interaction; WWP4.  
XX  
OS Synthetic.  
XX  
PN WO9737223-A1.  
XX  
PD 09-OCT-1997.  
XX  
PF 03-APR-1997; 97WO-US005547.  
XX  
PR 03-APR-1996; 96US-00630916.  
XX  
PA (CYTO-) CYTOGEN CORP.  
PA (UYNC-) UNIV NORTH CAROLINA.  
XX  
PI Pirozzi G, Kay BK, Fowlkes DM;  
XX  
DR WPI; 1997-503234/46.  
XX  
PT Identifying cell signalling and growth regulatory polypeptides by  
PT reaction with multivalent recognition complex - polypeptides are useful  
PT in targetted drug selection.  
XX  
PS Claim 49; Page 104; 220pp; English.  
XX  
CC Peptides AAW36798-801 and AAW37636-40 are peptides that contain WW  
CC domains. The WW domain is a small functional domain found in a large  
CC number of proteins from a variety of species including humans, nematodes  
CC and yeast. Its name is derived from the observation that two tryptophan  
CC residues, one in the amino terminal portion of the WW domain and one in  
CC the carboxyl terminal portion, are conserved. Most proteins containing WW  
CC domains have a function involving cell signalling and growth regulation  
CC or the organisation of the cytoskeleton. Polypeptides containing a WW  
CC domain are identified by treating a multivalent recognition unit complex  
CC with polypeptides and identifying those with selective affinity for the  
CC complex. Proteins containing WW domains are used for targeted drug  
CC screening, i.e. to identify potential modulators of specific WW domain  
CC interactions. The valency of the recognition unit is important in  
CC determining specificity of interaction with WW domains. In multivalent  
CC form specificity is relaxed, but not lost, so proteins containing WW  
CC domains similar, but not identical, to the sequence of the peptides,  
CC target WW can be detected, including new polypeptides  
XX  
SQ Sequence 38 AA;  
  
Query Match 71.3%; Score 92; DB 2; Length 38;  
Best Local Similarity 51.6%; Pred. No. 7.2e-09;  
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
  
QY 2 PLPGXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
| | | | | | | | | | | | | | | | | | | | | |

```
Db      2 PLPGWEKRTDPRGRFYVDHNTTTTQRP 32
RESULT 12
AAB83024
ID      AAB83024 standard; peptide; 38 AA.
XX
AC      AAB83024;
XX
XX      25-JUN-2001 (first entry)
XX
DE      Yeast Rsp5 HECT E3 ubiquitin ligase WW domain #3.
XX
XX      Yeast; HECT; homologous to E6 carboxyl terminus; ubiquitination;
KW      HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;
KW      antimicrobial; neuroprotective; transforming growth factor beta;
KW      TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;
KW      inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;
KW      renal failure; neurodegeneration; fibrosis; WW domain; Rsp5.
XX
XX      Saccharomyces cerevisiae.
OS
XX      WO200116604-A1.
XX
XX      08-MAR-2001.
XX
XX      29-AUG-2000; 2000WO-US023729.
XX
XX      30-AUG-1999; 99US-00385918.
XX
XX      (SIGN-) SIGNAL PHARM INC.
XX
XX      Hoekstra MF, Xie W, Murray BW, Mercurio FM;
PI
XX      WPI; 2001-327913/34.
XX
XX      Screening for modulators of TGF-beta and/or bone morphogenic protein
PT      (BMP) mediated signalling useful for treating cancer and osteoporosis by
PT      evaluating the ability of agents to modulate Smad protein degradation.
XX
XX      Disclosure; Page 14; 75pp; English.
XX
XX      The present sequence is the WW domain of a HECT (homologous to E6
CC      carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad
CC      PY motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.
CC      The sequence is provided in a specification relating to a method for
CC      screening for agents that modulate transforming growth factor (TGF)-beta
CC      and/or bone morphogenic protein (BMP)-mediated signalling. The method
CC      involves evaluating the effect of an agent on binding of HECT E3
CC      ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad
CC      protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein
CC      HECT E3 ubiquitin ligase activity. The method is useful for stimulating
CC      bone formation in a patient or treating a condition associated with
CC      insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that
CC      inhibit BMP-mediated signalling are useful for treating inflammation,
CC      ageing, cancer and infectious diseases. Agents that augment BMP-mediated
CC      signalling are useful for stimulating bone anabolism as well as treating
CC      broken bones, osteoporosis, and acute or chronic renal failure. Agents
CC      that inhibit TGF-mediated signalling are useful for treating cancer,
CC      inflammation, neurodegeneration and fibrosis
XX
XX      Sequence 38 AA;
Query Match      71.3%; Score 92; DB 4; Length 38;
Best Local Similarity 46.9%; Pred. No. 7.2e-09;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;
Oy      1 GPLPFGWEXXXXXXGXYYXXHNTTTTXXXP 32
      |||||
      : |||||
Db      1 GPLPFGWEMRLTNTARVYVDHNTKTTTWDPP 32
      |||||
      : |||||
RESULT 13
Query Match      71.3%; Score 92; DB 4; Length 38;
Best Local Similarity 46.9%; Pred. No. 7.2e-09;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;
Oy      1 GPLPFGWEXXXXXXGXYYXXHNTTTTXXXP 32
      |||||
      : |||||
Db      1 GPLPFGWEMRLTNTARVYVDHNTKTTTWDPP 32
      |||||
      : |||||
AAU87968
ID      AAU87968 standard; peptide; 38 AA.
XX
XX      AAU87968;
XX
XX      05-JUN-2002 (first entry)
XX
DE      Human WW domain #6.
XX
XX      Human; PDZ domain; WW domain; rat; cow; mouse; fruitfly; protein therapy;
KW      gene therapy; PDZ-mediated disease; inward potassium channel; WBP;
KW      dimer inhibitor peptide; carboxylate binding loop.
XX
XX      Homo sapiens.
OS
XX      WO200207751-A1.
XX
XX      31-JAN-2002.
XX
XX      24-JUL-2001; 2001WO-US023269.
XX
XX      25-JUL-2000; 2000US-0221215P.
XX
XX      28-NOV-2000; 2000US-00723810.
XX
XX      (AXCE-) AXCELL BIOSCIENCES CORP.
XX
XX      Herrero J, Pirozzi G, Uveges A;
PI
XX      WPI; 2002-195842/25.
XX
XX      Methods for identifying polypeptides comprising PDZ domains, the
PT      polypeptides and their encoding nucleic acids, useful for the diagnosis
PT      and treatment of PDZ related disorders.
XX
XX      Disclosure; Fig 20; 225pp; English.
XX
XX      The invention relates to methods for identifying polypeptides comprising
CC      PDZ domains, and their encoding nucleic acids. The sequences are used to
CC      identify modulators of their expression, function and activity, for use
CC      in the diagnosis and treatment of PDZ related disorders. Antibodies
CC      against the proteins and cells that produce them may be used for the
CC      treatment of PDZ-mediated disease states. Sequences AAU87843-AAU87974
CC      represent proteins containing PDZ domains, fragments of these proteins
CC      and other related peptides used in the methods of the invention
XX
XX      Sequence 38 AA;
Query Match      71.3%; Score 92; DB 5; Length 38;
Best Local Similarity 51.6%; Pred. No. 7.2e-09;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Oy      2 PLPFGWEXXXXXXGXYYXXHNTTTTXXXP 32
      |||||
      : |||||
Db      2 PLPFGWEKRTDPRGRFYVDHNTTTTQRP 32
      |||||
      : |||||
RESULT 14
AAB49229
ID      ADB49229 standard; peptide; 38 AA.
XX
XX      ADB49229;
XX
XX      04-DEC-2003 (first entry)
XX
XX      Novel human WWP2-2 protein WW domain.
DE
XX      WW domain; drug candidate screening; drug discovery; drug modification;
KW      drug refinement; immunogen; WW binding protein 2; WWP2; human.
XX
XX      Homo sapiens.
OS
XX      US2003077577-A1.
XX
```

```
PD 24-APR-2003.
XX
XX
XX 28-JUN-2002; 2002US-00185050.
XX
XX 03-APR-1996; 96US-00630916.
XX
XX 03-APR-1997; 97US-00826516.
XX
XX
XX (PIRO/) PIROZZI G.
XX (KAYB/) KAY B K.
XX (FOWL/) FOWLKES D M.
XX
XX Pirozzi G, Kay BK, Fowlkes DM;
XX
XX WPI; 2003-635075/60.
XX
XX Novel purified polypeptide comprising WW domain, useful for drug
XX discovery, modification and refinement, for discovering polypeptides
XX involved in pharmacological activities, or as an immunogen to generate
XX antibodies.
XX
XX Claim 98; Fig 5; 133pp; English.
XX
XX The invention describes a purified polypeptide (I) comprising a WW domain
XX which has a sequence (S1) selected from 11 sequences fully defined in the
XX specification, a sequence (S2) selected from 48 sequences fully defined
XX in the specification or a sequence (S3) comprising 683, 906, 224 or 725
XX amino acids fully defined in the specification. (I) is useful for
XX screening a potential drug candidate, by allowing (I) to come into
XX contact with at least one recognition unit having a selective affinity
XX for the WW domain in (I), in the presence of an amount of a potential
XX drug candidate, such that (I) and the recognition unit are capable of
XX interacting when brought into contact with one another in the absence of
XX the drug candidate, and determining the effect, if any, of the presence
XX of the amount of the drug candidate on the interaction of (I) with the
XX recognition unit. (I) is useful for drug discovery, modification and
XX refinement, for discovering polypeptides involved in pharmacological
XX activities, or as an immunogen to generate antibodies. This is the amino
XX acid sequence of novel human WW binding protein WWP2-2 WW domain.
XX
XX Sequence 38 AA;
XX
XX Query Match 71.3%; Score 92; DB 7; Length 38;
XX Best Local Similarity 51.6%; Pred. No. 7.2e-09;
XX Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
XX
XX 2 PLPGWEXXXXXXXGXXYYXXHNTXTTXXXP 32
XX ||||| ||||| : ||||| |||||
XX 2 PLPGWEKRTDPRGRFYVDHNTRTTTWQRP 32
XX
XX RESULT 15
XX ADB49220
XX ID ADB49220 standard; peptide; 38 AA.
XX
XX AC ADB49220;
XX
XX 04-DEC-2003 (first entry)
XX
XX Yeast WW binding protein Rsp5 #3.
XX
XX WW domain; drug candidate screening; drug discovery; drug modification;
XX drug refinement; immunogen; WW binding protein; yeast; Rsp5; WW domain.
XX
XX Ascomycota.
XX
XX US2003077577-A1.
XX
XX 24-APR-2003.
XX
XX 28-JUN-2002; 2002US-00185050.
XX
XX 03-APR-1996; 96US-00630916.
XX
XX 03-APR-1997; 97US-00826516.
XX
XX
XX (PIRO/) PIROZZI G.
XX (KAYB/) KAY B K.
XX (FOWL/) FOWLKES D M.
XX
XX Pirozzi G, Kay BK, Fowlkes DM;
XX
XX WPI; 2003-635075/60.
XX
XX Novel purified polypeptide comprising WW domain, useful for drug
XX discovery, modification and refinement, for discovering polypeptides
XX involved in pharmacological activities, or as an immunogen to generate
XX antibodies.
XX
XX Example; Fig 5; 133pp; English.
XX
XX The invention describes a purified polypeptide (I) comprising a WW domain
XX which has a sequence (S1) selected from 11 sequences fully defined in the
XX specification, a sequence (S2) selected from 48 sequences fully defined
XX in the specification or a sequence (S3) comprising 683, 906, 224 or 725
XX amino acids fully defined in the specification. (I) is useful for
XX screening a potential drug candidate, by allowing (I) to come into
XX contact with at least one recognition unit having a selective affinity
XX for the WW domain in (I), in the presence of an amount of a potential
XX drug candidate, such that (I) and the recognition unit are capable of
XX interacting when brought into contact with one another in the absence of
XX the drug candidate, and determining the effect, if any, of the presence
XX of the amount of the drug candidate on the interaction of (I) with the
XX recognition unit. (I) is useful for drug discovery, modification and
XX refinement, for discovering polypeptides involved in pharmacological
XX activities, or as an immunogen to generate antibodies. This is the amino
XX acid sequence of WW binding protein Rsp5 WW domain.
XX
XX Sequence 38 AA;
XX
XX Query Match 71.3%; Score 92; DB 7; Length 38;
XX Best Local Similarity 46.9%; Pred. No. 7.2e-09;
XX Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;
XX
XX 1 GPLPGWEXXXXXXXGXXYYXXHNTXTTXXXP 32
XX ||||| ||||| : ||||| |||||
XX 1 GPLPGWEMRLTWTARVYFVDHNTKTTTWD DP 32
XX
XX Search completed: October 13, 2005, 15:00:30
XX Job time : 75.3407 secs
```



OTHER INFORMATION: (e.g., S,H,P,D,E,T or Y)  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (16)...(16)  
OTHER INFORMATION: Xaa = independently selected polar amino acid  
OTHER INFORMATION: (e.g., S,H,P,D,E,T or Y)  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (17)...(17)  
OTHER INFORMATION: Xaa = any amino acid  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (20)...(20)  
OTHER INFORMATION: Xaa = hydrophobic residue (e.g., I,V,L or M)  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (21)...(21)  
OTHER INFORMATION: Xaa = any amino acid  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (25)...(25)  
OTHER INFORMATION: Xaa = independently selected polar amino acid  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (28)...(28)  
OTHER INFORMATION: Xaa = independently selected polar amino acid  
OTHER INFORMATION: (e.g., S,H,P,D,E,T or Y)  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (30)...(30)  
OTHER INFORMATION: Xaa = any amino acid  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (31)...(31)  
OTHER INFORMATION: Xaa = independently selected polar amino acid  
OTHER INFORMATION: (e.g., S,H,P,D,E,T or Y)  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (33)...(33)  
OTHER INFORMATION: Xaa = independently selected polar amino acid  
OTHER INFORMATION: (e.g., S,H,P,D,E,T or Y)  
US-10-307-956-2

Query Match 76.0%; Score 98; DB 14; Length 33;  
Best Local Similarity 100.0%; Pred. No. 4e-09;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXGXXYYXHNHTXTTXWXXP 32  
|||||  
Db 1 GPLPXGWEXXXXXXXGXXYYXHNHTXTTXWXXP 32

RESULT 2  
US-10-185-050-36  
Sequence 36, Application US/10185050  
Publication No. US2003007757A1  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
Kay, Brian K.  
Fowlkes, Dana M.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
NUMBER OF SEQUENCES: 233  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/185,050  
FILING DATE: 28-Jun-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,516  
FILING DATE: 03-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-208-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 896-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-10-185-050-36

Query Match 75.2%; Score 97; DB 14; Length 38;  
Best Local Similarity 53.1%; Pred. No. 6.7e-09;  
Matches 17; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXGXXYYXHNHTXTTXWXXP 32  
|||||  
Db 1 GPLPPGWKRQDVNGRVYYVHNHTRTTQWEDP 32

## RESULT 3

US-10-785-819-158  
Sequence 158, Application US/10785819  
Publication No. US20050112552A1  
GENERAL INFORMATION:  
APPLICANT: Herrero J  
APPLICANT: Pirozzi, G.  
APPLICANT: Uveges, A.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
POLYPEPTIDES HAVING PDZ DOMAINS AND METHODS OF USING SAME  
FILE REFERENCE: 1101-211  
CURRENT APPLICATION NUMBER: US/10/785,819  
CURRENT FILING DATE: 2004-02-23  
PRIOR APPLICATION NUMBER: US/09/723,810  
PRIOR FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 158  
LENGTH: 38  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-785-819-158

Query Match 75.2%; Score 97; DB 17; Length 38;  
Best Local Similarity 53.1%; Pred. No. 6.7e-09;  
Matches 17; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXXGXXYYXHNHTXTTXWXXP 32  
|||||  
Db 1 GPLPPGWKRQDVNGRVYYVHNHTRTTQWEDP 32

## RESULT 4

US-11-097-143-16857  
Sequence 16857, Application US/11097143  
Publication No. US20050208559A1



;; GENERAL INFORMATION:  
;; APPLICANT: Venter, J. Craig  
;; APPLICANT: et al.  
;; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
;; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
;; TITLE OF INVENTION: DROSOPHILA GENES.  
;; FILE REFERENCE: CL000728  
;; CURRENT APPLICATION NUMBER: US/11/097,143  
;; CURRENT FILING DATE: 2005-04-04  
;; PRIOR APPLICATION NUMBER: 60/157,832  
;; PRIOR FILING DATE: 1999-10-05  
;; PRIOR APPLICATION NUMBER: 60/160,191  
;; PRIOR FILING DATE: 1999-10-19  
;; PRIOR APPLICATION NUMBER: 60/161,932  
;; PRIOR FILING DATE: 1999-10-28  
;; PRIOR APPLICATION NUMBER: 60/164,769  
;; PRIOR FILING DATE: 1999-11-12  
;; PRIOR APPLICATION NUMBER: 60/173,383  
;; PRIOR FILING DATE: 1999-12-28  
;; PRIOR APPLICATION NUMBER: 60/175,693  
;; PRIOR FILING DATE: 2000-01-12  
;; PRIOR APPLICATION NUMBER: 60/184,831  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: 60/191,637  
;; PRIOR FILING DATE: 2000-03-23  
;; NUMBER OF SEQ ID NOS: 43008  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 16857  
;; LENGTH: 1082  
;; TYPE: PRT  
;; ORGANISM: DROSOPHILA  
US-11-097-143-16857

Query Match 73.6%; Score 95; DB 20; Length 1082;  
Best Local Similarity 50.0%; Pred. No. 3.5e-07;  
Matches 16; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXXYYXHXNTXTTXXXP 32  
||| ||| | : ||| |||  
Db 655 GPLPEGWEVRHTDGRVFYIDHNRTRTTQWEDP 686

RESULT 5  
US-10-185-050-39  
; Sequence 39, Application US/10185050  
; Publication No. US20030077577A1  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; Kay, Brian K.  
; Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 233  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/185,050  
; FILING DATE: 28-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/826,516  
; FILING DATE: 03-Apr-1997  
; ATTORNEY/AGENT INFORMATION:

;; NAME: MISROCK, S. LESLIE  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 1101-208-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 790-9090  
;; TELEFAX: (212) 896-8864/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 39:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 33 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: <Unknown>  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 5  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 12  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 13  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 14  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 16  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 20  
;; OTHER INFORMATION: /note= "A Polar Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 25  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 31  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 33  
;; OTHER INFORMATION: /note= "A Hydrophobic Amino Acid."  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-10-185-050-39

Query Match 72.1%; Score 93; DB 14; Length 33;  
Best Local Similarity 70.0%; Pred. No. 2.8e-08;  
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 LPXGWEXXXXXXXGXXYYXHXNTXTTXXXP 32  
||| ||| | : ||| |||  
Db 3 LPTGWEXXXTTTGTXYVYHXHNTTTTWTXTP 32

RESULT 6  
US-10-128-714-3162  
; Sequence 3162, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi



```
;
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
;
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-10-185-050-35

Query Match 71.3%; Score 92; DB 14; Length 38;
Best Local Similarity 51.6%; Pred. No. 4.7e-08;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
Db 2 PLPPGWEXRTDPRGRFYVDHNTTTTWRQP 32

RESULT 10
US-10-307-956-9
; Sequence 9, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-307-956-9

Query Match 71.3%; Score 92; DB 14; Length 38;
Best Local Similarity 46.9%; Pred. No. 4.7e-08;

;
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
;
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-10-185-050-35

Query Match 71.3%; Score 92; DB 14; Length 38;
Best Local Similarity 51.6%; Pred. No. 4.7e-08;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
Db 2 PLPPGWEXRTDPRGRFYVDHNTTTTWRQP 32

RESULT 11
US-10-785-819-164
; Sequence 164, Application US/10785819
; Publication No. US20050112552A1
; GENERAL INFORMATION:
; APPLICANT: Herrero, J.
; APPLICANT: Pirozzi, G.
; APPLICANT: Uveges, A.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING PDZ DOMAINS AND METHODS OF USING SAME
; FILE REFERENCE: 1101-211
; CURRENT APPLICATION NUMBER: US/10/785,819
; CURRENT FILING DATE: 2004-02-23
; PRIOR APPLICATION NUMBER: US/09/723,810
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-785-819-164

Query Match 71.3%; Score 92; DB 17; Length 38;
Best Local Similarity 51.6%; Pred. No. 4.7e-08;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
   |||||
Db 2 PLPPGWEXRTDPRGRFYVDHNTTTTWRQP 32

RESULT 12
US-10-313-955-4
; Sequence 4, Application US/10313955
; Publication No. US20030199038A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/313,955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-313-955-4

Query Match          71.3%; Score 92; DB 14; Length 766;
Best Local Similarity 46.9%; Pred. No. 8.1e-07;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPFGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
    ||||| ||| : ||||| |||
Db 345 GPLPFGWEMRLTNTARVYFVDHNTKTTTWD DP 376

RESULT 13
US-10-032-585-7296
; Sequence 7296, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-008-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7296
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7296

Query Match          71.3%; Score 92; DB 14; Length 832;
Best Local Similarity 46.9%; Pred. No. 8.7e-07;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPFGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
    ||||| ||| : ||||| |||
Db 410 GPLPFGWEMRLTNTARVYFVDHNTKTTTWD DP 441

RESULT 14
US-10-097-534-12
; Sequence 12, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: MATURATION
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-12

Query Match          71.3%; Score 92; DB 14; Length 870;
Best Local Similarity 51.6%; Pred. No. 9.1e-07;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPFGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
    ||||| ||| : ||||| |||
Db 331 PLPFGWEKRTDPRGRFYVDHNTTTTWQRP 361

RESULT 15
US-10-723-860-2167
; Sequence 2167, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882,0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2167
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-2167

Query Match          71.3%; Score 92; DB 16; Length 870;
Best Local Similarity 51.6%; Pred. No. 9.1e-07;
Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 PLPFGWEXXXXXXGXXYYXXHNTXTTXWXXP 32
    ||||| ||| : ||||| |||
Db 331 PLPFGWEKRTDPRGRFYVDHNTTTTWQRP 361

Search completed: October 13, 2005, 14:49:59
Job time : 71.3516 secs
```

GenCore version 5.1.6  
Copyright. (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:36:01 ; Search time 13.7802 Seconds  
(without alignments)  
230.414 Million cell updates/sec

Title: US-09-385-918-2  
Perfect score: 129  
Sequence: 1 GPLPFGWEXXXXXXXGXXYYXHNXTTXXWXXPX 33

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	72.1	786	2 T39585	ubiquitin protein
2	92	71.3	766	1 S66562	ubiquitin-protein
3	92	71.3	767	2 T37545	ubiquitin-protein
4	92	71.3	809	1 S43217	ubiquitin-protein
5	92	71.3	815	2 T49744	probable ubiquitin
6	91	70.5	1171	2 T42372	probable guanylate
7	90	69.8	472	2 B56954	yes-associated pro
8	87	67.4	383	2 JC7507	45K WW domain-cont
9	87	67.4	386	2 JC7508	45K WW domain-cont
10	87	67.4	820	2 T46412	ubiquitin-protein
11	86.5	67.1	1616	2 T47801	hypothetical prote
12	85	65.9	1277	2 T14152	synaptic scaffold
13	84	65.1	671	2 T37900	probable ubiquitin
14	83	64.3	457	2 G89632	protein F13E6.4 [i
15	83	64.3	708	2 I83196	NEDD-4 ORF - mouse
16	83	64.3	887	2 S70642	ubiquitin ligase N
17	81	62.8	1012	2 T23160	hypothetical prote
18	80	62.0	1256	2 J50209	brain-specific ang
19	79	61.2	454	2 A56954	yes-associated pro
20	75	58.1	448	2 I50730	yes-associated pro
21	68.5	53.1	616	2 T46292	hypothetical prote
22	66	51.2	691	2 T33637	hypothetical prote
23	58.5	45.3	163	2 S68520	peptidylprolyl iso
24	58.5	45.3	165	2 JC7136	Pin1 protein - hum
25	56.5	43.8	325	2 T02117	hypothetical prote
26	56	43.4	241	2 T25886	hypothetical prote
27	55	42.6	583	2 S30014	hypothetical prote
28	55	42.6	870	2 A41130	dystrophin homolo
29	55	42.6	3660	1 S02041	dystrophin, muscle

30	55	42.6	3678	2 S28916	dystrophin - mouse
31	55	42.6	3685	1 A27605	dystrophin, muscle
32	52	40.3	987	2 I48373	G-utrophin - mouse
33	52	40.3	3433	1 S28381	utrophin - human
34	51.5	39.9	166	2 T08426	pin1 protein homol
35	51.5	39.9	499	2 S22571	integrase-like pro
36	51	39.5	158	2 C70976	hypothetical prote
37	48.5	37.6	234	2 T49448	hypothetical prote
38	48	37.2	465	2 S69038	hypothetical prote
39	47.5	36.8	397	2 T09579	nuclear protein EM
40	47	36.4	120	2 T26779	hypothetical prote
41	45	34.9	311	2 JN0520	beta-lactamase [EC
42	45	34.9	1590	2 B87754	protein C43E11.3 [
43	44	34.1	26	2 S64716	formin binding pro
44	44	34.1	411	2 S58094	hypothetical WW do
45	44	34.1	1098	2 T08599	probable transcrip

ALIGNMENTS

RESULT 1

T39585  
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T39585

R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1997

A:Reference number: 221865

A:Accession: T39585

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-786 <VOL>

A:Cross-references: UNIPROT:O14326; EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GN00067; SPDB:SF

A:Experimental source: strain 972h-; cosmid cl6E9

C:Genetics:

A:Gene: SPDB:SPBC16E9.1lc

A:Map position: 2

A:Introns: 60/2

C:Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology

F:236-273/Domain: WW repeat homology <WW1>

F:306-343/Domain: WW repeat homology <WWR2>

F:364-401/Domain: WW repeat homology <WWR3>

Query Match 72.1%; Score 93; DB 2; Length 786;

Best Local Similarity 50.0%; Pred. No. 1.8e-08;

Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPFGWEXXXXXXXGXXYYXHNXTTXXWXXP 32

Db 236 GRLPFGWERRADSLGRTYYVDHNRTRTTWTRP 267

RESULT 2

S66562

ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: E6-AP-like protein ubiquitin ligase

C:Species: Schizosaccharomyces pombe

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S66562; T45159

R:Nefsky, B.; Beach, D.

EMBO J. 15, 1301-1312, 1996

A:Title: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25A

A:Reference number: S66562; MUID:96205868; PMID:8635463

A:Accession: S66562

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-766 <NEF>

A:Cross-references: UNIPROT:Q92462; GB:U66716; NID:g1519443; PIDN:AAB07514.1; PID:g15194

R:Nefsky, B.S.; Beach, D.

submitted to the EMBL Data Library, August 1996

A:Description: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of

A:Reference number: Z22935  
A:Accession: T45159  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-766 <N52>  
A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1  
C:Genetics:  
A:Gene: publ  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C:Keywords: ligase  
F:205-242/Domain: WW repeat homology <WW1>  
F:288-325/Domain: WW repeat homology <WW2>  
F:345-382/Domain: WW repeat homology <WW3>

Query Match 71.3%; Score 92; DB 1; Length 766;  
Best Local Similarity 46.9%; Pred. No. 2.6e-08;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
||||| ||| : ||| ||| |  
Db 345 GPLPSGWEMLTNTARVYFVDHNTKTTTWDPP 376

RESULT 3  
T37545  
ubiquitin-protein ligase (EC 6.3.2.19) publ [validated] - fission yeast (Schizosaccharom  
N:Alternate names: ubiquitin ligase Publ  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
A:Accession: T37545; T48655  
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z21722  
A:Accession: T37545  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-767 <MCI>  
A:Cross-references: UNIPROT:Q92462; EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:SF  
R:Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.  
Mol. Gen. Genet. 254, 520-528, 1997  
A:Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning publ ub  
A:Reference number: Z07985; MUID:97340937; PMID:9197411  
A:Accession: T48655  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-767 <SAL>  
A:Cross-references: EMBL:U62795; NID:G2262192; PIDN:AAB63350.1; PID:G2262193  
A:Experimental source: strain J227  
C:Genetics:  
A:Gene: publ; SPDB:SPAC11G7.02  
A:Map position: 1  
A:Introns: 6/2; 14/1; 62/2  
C:Function: <UBI>  
C:Function: <CYC>  
A:Description: involved in of the mitotic activating tyrosine phosphatase cdc25 [validat  
C:Function: <TOL>  
A:Description: required for low pH-tolerance [validated, MUID:97340937]  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C:Keywords: cell cycle control; ligase  
F:205-242/Domain: WW repeat homology <WW1>  
F:288-325/Domain: WW repeat homology <WW2>  
F:345-382/Domain: WW repeat homology <WW3>

Query Match 71.3%; Score 92; DB 2; Length 767;  
Best Local Similarity 46.9%; Pred. No. 2.6e-08;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
||||| ||| : ||| ||| |  
Db 345 GPLPSGWEMLTNTARVYFVDHNTKTTTWDPP 376

RESULT 4  
S43217  
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein  
C:Species: Saccharomyces cerevisiae  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
A:Accession: S43217; S50628; S70050  
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S30812  
A:Accession: S43217  
A:Molecule type: DNA  
A:Residues: 1-809 <MUL>  
A:Cross-references: UNIPROT:P39940; GB:U18916; EMBL:L11119; NID:gl384128; PIDN:AAC03223;  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda  
A:Reference number: S50628  
A:Accession: S50628  
A:Molecule type: DNA  
A:Residues: 1-809 <DIE>  
A:Cross-references: EMBL:U18916; NID:gl384128; PIDN:AAC03223.1; PID:9603364; GSPDB:GN000  
R:Hein, C.; Springael, J.Y.; Volland, C.; Hagenauer-Tsapis, R.; Andre, B.  
Mol. Microbiol. 18, 77-87, 1995  
A:Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur4  
A:Reference number: S70050; MUID:96154942; PMID:8596462  
A:Accession: S70050  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-101 <HEI>  
C:Genetics:  
A:Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125w  
A:Cross-references: MIPS:YER125w; SGD:S0000927  
A:Map position: 5R  
C:Function:  
A:Description: involved in endocytosis of Gap1 protein and FUR4 protein; binds and ubiq  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C:Keywords: ligase  
F:229-266/Domain: WW repeat homology <WW1>  
F:331-368/Domain: WW repeat homology <WW2>  
F:387-424/Domain: WW repeat homology <WW3>

Query Match 71.3%; Score 92; DB 1; Length 809;  
Best Local Similarity 46.9%; Pred. No. 2.8e-08;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPXGWEXXXXXXGXXYYXXHNTXTTXWXXP 32  
||||| ||| : ||| ||| |  
Db 387 GPLPSGWEMLTNTARVYFVDHNTKTTTWDPP 418

RESULT 5  
T49744  
probable ubiquitin-protein ligase [imported] - Neurospora crassa  
N:Alternate names: protein B24B19.160  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
A:Accession: T49744  
R:Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49744  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-815 <SCH>  
A:Cross-references: EMBL:ALJ56192; GSPDB:GN00116; NCSP:B24B19.160  
A:Experimental source: BAC clone B24B19; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B24B19.160  
A:Map position: 6  
A:Introns: 11/1; 24/1; 59/2; 110/1; 783/2  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology



```

F:348-395/Domain: WW repeat homology <WW2>

Query Match      65.9%; Score 85; DB 2; Length 1277;
Best Local Similarity 45.2%; Pred. No. 7.9e-07;
Matches 14; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

QY  2  PLPXGWEHXXXXXGXYYXXHNTXTTXWXXP 32
      ||| || | | : | || | | | |
DB  303  PLPDNEMAYTERKEGVYFIDHNTKTTTSLDP 333

RESULT 13
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37900
R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Accession: T37900
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <RIE>
A:Cross-references: UNIPROT:Q9UTG2; EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN0000
A:Experimental source: strain 972H-; cosmid c1805
C:Genetics:
A:Gene: SPDB:SPAC1805.15c
A:Map position: 1
A:Map position: 60/2; 105/1; 639/2
A:Introns: 60/2; 105/1; 639/2
F:242-279/Domain: WW repeat homology <WWR>

Query Match      65.1%; Score 84; DB 2; Length 671;
Best Local Similarity 43.8%; Pred. No. 5.9e-07;
Matches 14; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

QY  1  GPLPXGWEHXXXXXGXYYXXHNTXTTXWXXP 32
      ||| || | | : | || | | | |
DB  242  GPLPAGWEMRLSEYHYVYFVDHSTKTTTWSDP 273

RESULT 14
G89632
protein F13E6.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: G89632
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigation
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Project
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103,
A:Accession: G89632
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: UNIPROT:Q19404; GB:chr_X; PIDN:CAA92121.1; PID:g3875841; GSPDB:
C:Genetics:
A:Gene: F13E6.4
A:Map position: X

Query Match      64.3%; Score 83; DB 2; Length 457;
Best Local Similarity 41.9%; Pred. No. 5.9e-07;
Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY  2  PLPXGWEHXXXXXGXYYXXHNTXTTXWXXP 32
      ||| || | | : | || | | | |
DB  219  PMQGWEMCYDSGVRYFKDHNSKTTTWDPP 249

RESULT 15
I183196

```



NEDD-4 ORF - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 18-Aug-2000  
 C:Accession: I83196  
 R:Kumar, S.; Tomooka, Y.; Noda, M.  
 Biochem. Biophys. Res. Commun. 185, 1155-1161, 1992  
 A:Title: Identification of a set of genes with developmentally down-regulated expression  
 A:Reference number: I60167; MUID:92328780; PMID:1378265  
 A:Accession: I83196  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <RES>  
 A:Cross-references: GB:D10714; NID:g220508; PID:g220509  
 C:Genetics:  
 A:Gene: NEDD-4  
 C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin  
 F:40-77/Domain: WW repeat homology <WW1>  
 F:196-233/Domain: WW repeat homology <WW2>  
 F:251-288/Domain: WW repeat homology <WW3>  
 F:347-682/Domain: ubiquitin-protein ligase homology <UBI>  
  
 Query Match 64.3%; Score 83; DB 2; Length 708;  
 Best Local Similarity 46.7%; Pred. No. 9.5e-07;  
 Matches 14; Conservative 1; Mismatches 15; Indels 0; Gaps 0;  
  
 Qy 3 LPXGWEXXXXXGXXYYXHNXTTXXXP 32  
 ||||| | ||||| |  
 Db 198 LPPGWEEXQDDRGSRYYVDHNSKTTWSKP 227

Search completed: October 13, 2005, 15:11:14  
 Job time : 15.7802 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 62.8571 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-2  
Perfect score: 129  
Sequence: 1 GPLPQGWEXXXXXXXHNTTXXXPX 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_eprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	95	73.6	793	Q7KUR2	Q7kur2 drosophila
2	95	73.6	834	Q95R64	Q95r64 drosophila
3	95	73.6	838	Q95TQ0	Q95tq0 drosophila
4	95	73.6	956	Q81QR6	Q81qr6 drosophila
5	95	73.6	1007	Q9VW13	Q9rv13 drosophila
6	93	72.1	786	PUB3_SCHPO	Ol4326 schizosacch
7	93	72.1	945	Q7PW66	Q7pw66 anopheles g
8	92	71.3	767	PUB1_SCHPO	Q92462 schizosacch
9	92	71.3	781	Q6BT41	Q6bt41 debaryomyce
10	92	71.3	787	Q6M906	Q6m906 neurospora
11	92	71.3	792	Q95XU3	Q95xu3 caenorhabdi
12	92	71.3	794	Q9N2Z7	Q9n2z7 caenorhabdi
13	92	71.3	806	Q7RV01	Q7rv01 neurospora
14	92	71.3	809	R3P5_YEAST	P39940 saccharomyc
15	92	71.3	817	Q75AI2	Q75ai2 ashbya goss
16	92	71.3	819	Q6CNC7	Q6cnc7 kluyveromyc
17	92	71.3	822	Q6FN71	Q6fn71 candida gla
18	92	71.3	854	Q6C5H0	Q6c5h0 yarrowia li
19	92	71.3	870	WMP2_HUMAN	O00308 homo sapien
20	92	71.3	870	WMP2_MOUSE	Q9dbh0 mus musculu
21	92	71.3	918	WMP1_MOUSE	Q8bzz3 mus musculu
22	92	71.3	1048	Q7POR5	Q7por5 anopheles g
23	91	70.5	1471	BAP1_MOUSE	Q6rhr9 mus musculu
24	91	70.5	1491	BAP1_HUMAN	Q96qz7 homo sapien
25	90	69.8	397	Q86T74	Q86t74 homo sapien
26	90	69.8	472	YAP1_MOUSE	P46938 mus musculu
27	90	69.8	488	Q7Z574	Q7z574 homo sapien
28	90	69.8	488	Q91WL1	Q91wl1 mus musculu
29	90	69.8	504	Q81UY9	Q81uy9 homo sapien
30	89	69.0	806	Q6P289	Q6p289 mus musculu
31	89	69.0	864	ITCH_MOUSE	Q8c863 mus musculu

32	89	69.0	903	1	ITCH_HUMAN	Q96j02 homo sapien
33	89	69.0	949	2	Q9Y0H4	Q9y0h4 drosophila
34	88	68.2	922	1	WMP1_HUMAN	Q9h0m0 homo sapien
35	88	68.2	970	2	Q6DIR6	Q6dir6 xenopus tro
36	88	68.2	971	2	O42573	O42573 xenopus lae
37	88	68.2	971	2	Q6GMD5	Q6gmd5 xenopus lae
38	87	67.4	138	2	Q6H2Y6	Q6h2y6 trypanosoma
39	87	67.4	383	1	SAV1_HUMAN	Q9h4b6 homo sapien
40	87	67.4	383	2	Q6IA58	Q6ia58 homo sapien
41	87	67.4	386	1	SAV1_MOUSE	Q8veb2 mus musculu
42	87	67.4	571	2	Q8QGJ2	Q8qgj2 gallus gall
43	87	67.4	820	2	Q9NT88	Q9nt88 homo sapien
44	87	67.4	854	2	Q9H2W4	Q9h2w4 homo sapien
45	87	67.4	855	2	Q9BW58	Q9bw58 homo sapien

ALIGNMENTS

RESULT 1  
Q7KUR2  
ID Q7KUR2 PRELIMINARY; PRT; 793 AA.  
AC Q7KUR2;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE CG7555-PD.  
GN Name=Nedd4; ORFName=CG7555;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.B.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."

Science 287:2185-2195(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
 George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 R Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminer J.S., Bergman C.M., Krommiller B., Carlson J., Svirska R.,  
 Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomics perspective."  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 Hradecky P., Huang Y., Kaminer J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review."  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 C2 domain.  
 DR EMBL; AE003523; AAS64973.1; -.  
 DR HSP; P04410, 1A25.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2 CalB.  
 DR InterPro; IPR000569; HECT.  
 DR InterPro; IPR001202; WW Rep5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 3.  
 DR PRINTS; PR00360; C2DOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTc; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS00499; C2 DOMAIN 1; UNKNOWN\_1.  
 DR PROSITE; PS00004; C2 DOMAIN\_2; 1.  
 DR PROSITE; PS02037; HECT; 1.  
 DR PROSITE; PS01159; WW DOMAIN 1; 3.  
 DR PROSITE; PS00020; WW DOMAIN\_2; 3.  
 SQ SEQUENCE 793 AA; 91167 MW; 56D9B230DCE32F4F CRC64;

Query Match 73.6%; Score 95; DB 2; Length 793;

Best Local Similarity 50.0%; Pred. No. 2.6e-08;

Matches 16; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GPLPQGWEXXXXXXXHNTTXXXP 32  
 ||||| |||| | : | ||| | | |

366 GPLPEGWEERVHTDGRVFIIDHNTRTTQWEDP 397  
 Db  
 RESULT 2  
 Q95R64 PRELIMINARY; PRT; 834 AA.  
 AC Q95R64;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE SD04682p (CG7555-PC).  
 GN Name=Nedd4; ORFNames=CG7555;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacle J., Faragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Batwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Massarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang W., Zhou X., Zhu S., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirska R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupoy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 C2 domain.  
DR EMBL; AY061595; AAL29143.1; -.  
DR EMBL; AE003523; AAN11694.1; -.  
DR HSP; Q62940; 115H.  
DR FlyBase; FBgn0036736; Nedda4.  
DR GO; GO:0016567; P:protein ubiquitination; NAS.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR PRINTS; PR00360; C2DOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS00499; C2 DOMAIN 1; UNKNOWN\_1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW DOMAIN 1; 3.  
DR PROSITE; PS00020; WW DOMAIN 2; 3.  
SQ SEQUENCE 834 AA; 95302 MW; CD0B385C93C95B00 CRC64;  
  
Query Match 73.6%; Score 95; DB 2; Length 834;  
Best Local Similarity 50.0%; Pred. No. 2.8e-08;  
Matches 16; Conservative 1; Mismatches 15; Indels 0; Gaps 0;  
  
QY 1 GPLPVGWEXXXXXXGXXVYHNTTXXXP 32  
DB 407 GPLPVGWEEVHTDGRVFYDHTTRTTQWEDP 438  
  
RESULT 3  
Q95TQ0  
ID Q95TQ0 PRELIMINARY; PRT; 838 AA.  
AC Q95TQ0;  
DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE LD1220p.  
GN Name=Nedd4;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkelley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY058619; AAL13848.1; -.  
DR HSP; Q62940; 115H.  
DR FlyBase; FBgn0036736; Nedda4.  
DR GO; GO:0016567; P:protein ubiquitination; NAS.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 3.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW DOMAIN 1; 3.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 3.  
SQ SEQUENCE 838 AA; 96400 MW; A3997EA68B4A08D0 CRC64;  
  
Query Match 73.6%; Score 95; DB 2; Length 838;  
Best Local Similarity 50.0%; Pred. No. 2.8e-08;  
Matches 16; Conservative 1; Mismatches 15; Indels 0; Gaps 0;  
  
QY 1 GPLPVGWEXXXXXXGXXVYHNTTXXXP 32  
DB 411 GPLPVGWEEVHTDGRVFYDHTTRTTQWEDP 442  
  
RESULT 4  
Q91QR6  
ID Q91QR6 PRELIMINARY; PRT; 956 AA.  
AC Q91QR6;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE CG7555-PB.  
GN Name=Nedd4; ORFNames=CG7555;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,





CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: Contains 3 WW domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z99759; CAB16903.1; -;  
 CC EMBL; AB001023; BAA19217.1; -;  
 CC PIR; T39585; T39585.  
 CC HSP; O13526; 1PIN.  
 CC GeneDB Spombe; SPBC16E9.11c; -;  
 CC InterPro; IPR000008; C2; CaLB.  
 CC InterPro; IPR008973; C2; CaLB.  
 CC InterPro; IPR000569; HECT.  
 CC InterPro; IPR002349; WW.  
 CC InterPro; IPR001202; WW\_Rep5\_WWP.  
 CC Pfam; PF00168; C2; 1.  
 CC Pfam; PF00632; HECT; 1.  
 CC Pfam; PF00397; WW; 3.  
 CC PRINTS; PR00403; WWDOMAIN.  
 CC SMART; SM00239; C2; 1.  
 CC SMART; SM00119; HECTc; 1.  
 CC SMART; SM00456; WW; 3.  
 CC PROSITE; PS00499; C2 DOMAIN\_1; 1.  
 CC PROSITE; PS00004; C2 DOMAIN\_2; 1.  
 CC PROSITE; PS0237; HECT; 1.  
 CC PROSITE; PS01159; WW DOMAIN\_1; 3.  
 CC PROSITE; PS0020; WW DOMAIN\_2; 3.  
 CC Ligase; Repeat; Ubl conjugation; Ubl conjugation pathway.  
 KW DOMAIN 1 90  
 FT DOMAIN 144 217 Thr-rich.  
 FT DOMAIN 236 269 WW 1.  
 FT DOMAIN 306 339 WW 2.  
 FT DOMAIN 364 397 WW 3.  
 FT DOMAIN 453 786 HECT.  
 FT BINDING 754 754 Ubiquitin (By similarity).  
 FT CONFLICT 637 649 FNAYSGFVELVS -> LMHFLVLLNWP (in Ref.  
 FT 2).  
 SQ SEQUENCE 786 AA; 89259 MW; 57B7A859F5497B9A CRC64;  
 Query Match 72.1%; Score 93; DB 1; Length 786;  
 Best Local Similarity 50.0%; Pred. No. 5.9e-08;  
 Matches 16; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 GPLPFGWEXXXXXXXGXXVXXHNTXTTXWXXP 32  
 DB 236 GRLPFGWERRADSLGRYYVVDHNTRTTWRP 267  
 RESULT 7  
 ID Q7PW66 PRELIMINARY; PRT; 945 AA.  
 AC Q7PW66;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)  
 DE ENSANGP0000016497 (Fragment).  
 GN Name=ENSANGG0000014008;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008984; EAA14748.1; -;  
 DR HSP; Q9H0M0; IND7.  
 DR GO; GO:0005222; C:intracellular; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2 CaLB.  
 DR InterPro; IPR000569; HECT.  
 DR InterPro; IPR001202; WW\_Rep5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 4.  
 DR PROSITE; PS00004; C2 DOMAIN\_2; 1.  
 DR PROSITE; PS0237; HECT; 1.  
 DR PROSITE; PS01159; WW DOMAIN\_1; 3.  
 DR PROSITE; PS0020; WW DOMAIN\_2; 2.  
 FT NON TER 1  
 SQ SEQUENCE 945 AA; 106764 MW; 5AE9827A2748BB11 CRC64;  
 Query Match 72.1%; Score 93; DB 2; Length 945;  
 Best Local Similarity 51.6%; Pred. No. 7.1e-08;  
 Matches 16; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 2 PLPFGWEXXXXXXXGXXVXXHNTXTTXWXXP 32  
 DB 394 PLPAGWEQRDPGRVYVVDHNTRTTWRP 424  
 RESULT 8  
 ID PUB1 SCHPO STANDARD; PRT; 767 AA.  
 AC Q92462; O14454;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-DEC-1998 (Rel. 37, last sequence update)  
 DT 25-OCT-2004 (Rel. 45, last annotation update)  
 DE E3 ubiquitin-protein ligase publi (BC 6.3.2.-).  
 GN Name=pubi; ORFNames=SPAC11G7.02;  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96205868; PubMed=8635463;  
 RA Nefsky B., Beach D.;  
 RT "Pubi acts as an EG-AP-like protein ubiquitin ligase in the  
 RT degradation of cdc25.";  
 RL EMBO J. 15:1301-1312(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J227;  
 RC MEDLINE=97340937; PubMed=9197411; DOI=10.1007/s004380050447;  
 RA Saleki R., Jia Z., Karagiannis J., Young P.G.;  
 RT "Tolerance of low pH in Schizosaccharomyces pombe requires a  
 RT functioning publ ubiquitin ligase";  
 RL Mol. Gen. Genet. 254:520-528(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,



RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jegelke K.,  
RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.-J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas R.R., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
RA "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
RN [4]  
RP SUBCELLULAR LOCATION.  
RX PubMed=11956316;  
RA Tamai K.K., Shimoda C.;  
RT "The novel HECT-type ubiquitin-protein ligase Pub2p shares partially  
RT overlapping function with Pub1p in Schizosaccharomyces pombe.";  
RL J. Cell Sci. 115:1847-1857(2002).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
CC then directly transfers the ubiquitin to targeted substrates.  
CC Regulates ubiquitination of cdc25.  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBCELLULAR LOCATION: Membrane-associated, and cytoplasmic.  
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
CC thioester formation.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -!- SIMILARITY: Contains 3 WW domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y07592; CAA68867.1; -;  
DR EMBL; U66716; AAB07514.1; -;  
DR EMBL; Z99161; CAB16207.1; -;  
DR EMBL; U62795; AAB63350.1; -;  
DR PIR; S66562; S66562.  
DR PIR; T37545; T37545.  
DR HSSP; Q62940; 115H.  
DR GeneDB SPombe; SPAC11G7.02; -;  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WMP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 3.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS00499; C2 DOMAIN 1; 1.  
DR PROSITE; PS50004; C2 DOMAIN 2; 1.  
DR PROSITE; PS50237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN 1; 3.  
DR PROSITE; PS50020; WW DOMAIN 2; 3.  
KW Ligase; Repeat; Ub1 conjugation; Ub1 conjugation pathway.  
FT DOMAIN 17 32 C2 domain.  
FT DOMAIN 211 236 WW 1.  
FT DOMAIN 242 247 Poly-Ala.  
FT DOMAIN 294 319 WW 2.  
FT DOMAIN 351 376 WW 3.  
FT DOMAIN 735 767 HECT.  
FT BINDING 735 735 Ubiquitin (By similarity).  
FT CONFLICT 163 163 Q -> K (in Ref. 1).  
FT CONFLICT 609 609 Missing (in Ref. 1).  
FT CONFLICT 661 661 T -> K (in Ref. 1).  
SQ SEQUENCE 767 AA; 87267 MW; F1455A155EB9ACF7 CRC64;  
Query Match 71.3%; Score 92; DB 1; Length 767;  
Best Local Similarity 46.9%; Pred. No. 8.8e-08;  
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;  
QY 1 GPLPQGWEXXXXXXXGXXYYXXHNTYTXWXXP 32  
| | | | | : | | | | | |  
DB 345 GPLPQGWEMRLTNTARVYFVDHNTKTTTWD 376  
RESULT 9  
Q6BT41 PRELIMINARY; PRT; 781 AA.  
AC Q6BT41;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Similar to sp|P39940 Saccharomyces cerevisiae Ubiqutin--protein  
DE ligase RSP5.  
DE ORFNames=DEHA0D0422g;  
GN Debaryomyces hansenii CBS767.  
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.  
OX NCBI\_TaxID=284592;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS767;  
RG Genolevures;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talia E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpetti C., Gaillardin C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT "Genome evolution in yeasts.";  
RL Nature 430:35-44(2004).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS767;  
RG Genoscope;  
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
DR EMBL; CR382136; CAG86767.1; -;  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR001202; WW\_Rsp5\_WMP.

```

DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTc; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00004; C2 DOMAIN 2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 3.
DR PROSITE; PS50020; WW DOMAIN 2; 3.
KW Ligase.
SQ SEQUENCE 781 AA; 89368 MW; 98FA2CAAFD92754 CRC64;

Query Match 71.3%; Score 92; DB 2; Length 781;
Best Local Similarity 46.9%; Pred. No. 8.9e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPFGWEXXXXXXXGXXHNTXTTXXXP 32
   |||||
   : : : : :
Db 359 GPLPFGWEMRLTNTARVYFVDHNTKTTTWD 390

RESULT 10
Q6M906 PRELIMINARY; PRT; 787 AA.
ID Q6M906
AC Q6M906;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Probable ubiquitin-protein ligase.
GN Name=G17A4.220;
OS Neosporea crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 C2 domain.
DR EMBL; EX008812; CAP06157.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTc; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2 DOMAIN 1; UNKNOWN_1.
DR PROSITE; PS50004; C2 DOMAIN 2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 3.
DR PROSITE; PS50020; WW DOMAIN 2; 3.
KW Ligase.
SQ SEQUENCE 787 AA; 89265 MW; 14E9392E5AA3B923 CRC64;

Query Match 71.3%; Score 92; DB 2; Length 787;
Best Local Similarity 46.9%; Pred. No. 9e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GPLPFGWEXXXXXXXGXXHNTXTTXXXP 32
   |||||
   : : : : :
Db 365 GPLPFGWEMRLTNTARVYFVDHNTKTTTWD 396

RESULT 11
Q95XU3 PRELIMINARY; PRT; 792 AA.
ID Q95XU3
AC Q95XU3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ww domain protein (E3 ubiquitin ligase) protein 1, isoform b.
GN Name=wwp-1; ORFNames=Y65B4BR.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Maupin R., Bauer C.;
RT "The sequence of C. elegans cosmid Y65B4BR.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[7]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[8]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
[9]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
[10]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

```





```

CC      on RBPI.
CC      -!- PATHWAY: Ubiquitin conjugation; third step.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
CC      -!- PTM: The ubiquitination appears to be the result of an
CC      intramolecular transfer of ubiquitin.
CC      -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC      thiolester formation.
CC      -!- SIMILARITY: Contains 1 C2 domain.
CC      -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC      domain.
CC      -!- SIMILARITY: Contains 3 WW domains.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U18916; AAC03223.1; -.
DR      PIR; S43217; S43217.
DR      HSSP; Q62940; IISH.
DR      IntAct; P39940; -.
DR      GeneOnline; I39204; -.
DR      SGD; S000000527; RSP5.
DR      GO; GO:0000151; C:ubiquitin ligase complex; IDA.
DR      GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR      GO; GO:0006333; P:chromatin assembly/disassembly; IMP.
DR      GO; GO:0006513; P:protein monoubiquitination; IDA.
DR      GO; GO:0000209; P:protein polyubiquitination; IDA.
DR      InterPro; IPR000008; C2.
DR      InterPro; IPR008973; C2_CalB.
DR      InterPro; IPR000569; HECT.
DR      InterPro; IPR002349; WW.
DR      InterPro; IPR001202; WW_Rsp5_WWP.
DR      Pfam; PF00168; C2; 1.
DR      Pfam; PF00632; HECT; 1.
DR      PRINTS; PR00397; WW; 3.
DR      SMART; SM00119; HECTC; 1.
DR      SMART; SM00456; WW; 3.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS50004; C2_DOMAIN_2; 1.
DR      PROSITE; PS50237; HECT; 1.
DR      PROSITE; PS01159; WW_DOMAIN_1; 3.
DR      PROSITE; PS50020; WW_DOMAIN_2; 3.
KW      Ligase; Repeat; Ubl conjugation; Ubl conjugation pathway.
FT      DOMAIN 1 88 C2 domain.
FT      DOMAIN 229 262 WW 1.
FT      DOMAIN 315 322 Poly-Ala.
FT      DOMAIN 331 364 WW 2.
FT      DOMAIN 387 420 WW 3.
FT      DOMAIN 705 809 HECT.
FT      BINDING 777 777 Ubiquitin.
FT      VARIANT 733 733 L -> S (in RSP5-1; impaired in ubiquitin-
FT      thioester formation and catalysis of
FT      substrate ubiquitination).
FT      MUTAGEN 777 777 C->A: Loss of ubiquitination.
SQ      SEQUENCE 809 AA; 91816 MW; 6F1836384479E70F CRC64;

Query Match 71.3%; Score 92; DB 1; Length 809;
Best Local Similarity 46.9%; Pred. No. 9.3e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy      1 GPLPQGWEXXXXXXXGXXYYXHXNTXTTXXXP 32
Db      387 GPLPSGWMRLTNTARVYFVDHNTKTTTWDPP 418

RESULT 15
Q75A12

```

```

ID      Q75A12 PRELIMINARY; PRT; 817 AA.
AC      Q75A12;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      ADL055CP.
GN      ORFNames=ADL055C;
OS      Asbya gossypii (Yeast) (Eremothecium gossypii).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX      NCBI_TaxID=33169;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 10895;
RA      Voegelé S.E., Brachat S., Dietrich F.S., Lerch A., Galfney T.,
RA      Philippsen P.;
RL      Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: Contains 1 C2 domain.
CC      EMBL; AE016817; AAS51865.1; -.
DR      HSSP; P47712; IBC1.
DR      AGD; ADL055C; -.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR      GO; GO:0006512; P:ubiquitin cycle; IEA.
DR      InterPro; IPR000008; C2_CalB.
DR      InterPro; IPR008973; C2_CalB.
DR      InterPro; IPR000569; HECT.
DR      InterPro; IPR002349; WW.
DR      InterPro; IPR001202; WW_Rsp5_WWP.
DR      Pfam; PF00168; C2; 1.
DR      Pfam; PF00632; HECT; 1.
DR      PRINTS; PR00397; WW; 3.
DR      SMART; SM00239; C2; 1.
DR      SMART; SM00119; HECTC; 1.
DR      SMART; SM00456; WW; 3.
DR      PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR      PROSITE; PS50004; C2_DOMAIN_2; 1.
DR      PROSITE; PS50237; HECT; 1.
DR      PROSITE; PS01159; WW_DOMAIN_1; 3.
DR      PROSITE; PS50020; WW_DOMAIN_2; 3.
SQ      SEQUENCE 817 AA; 51996 MW; 52FD02B4FB9C48E8 CRC64;

Query Match 71.3%; Score 92; DB 2; Length 817;
Best Local Similarity 46.9%; Pred. No. 9.3e-08;
Matches 15; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy      1 GPLPQGWEXXXXXXXGXXYYXHXNTXTTXXXP 32
Db      395 GPLPSGWMRLTNTARVYFVDHNTKTTTWDPP 426

Search completed: October 13, 2005, 15:09:15
Job time : 65.8571 secs

```

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 85.6044 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-3

Perfect score: 218

Sequence: 1 SPLPPGWEERDILGRYYVNHESRRTQWKRPDPQDNL 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218	100.0	38	4	AAB83018 Human Ned
2	218	100.0	733	6	Aae32726 Human Ned
3	218	100.0	900	8	Adq07991 Human Ned
4	218	100.0	927	2	Aay25170 Human KIA
5	218	100.0	927	6	Aae32725 Human Ned
6	209	95.9	38	4	AAB83019 Human Ned
7	209	95.9	38	7	Adb49212 Mouse WW
8	203	93.1	38	2	Aar97680 Mouse Ned
9	203	93.1	38	3	AAB21978 Nedda4/mou
10	203	93.1	38	7	Adb49211 Human WW
11	148	67.9	818	2	Aaw13386 Human pro
12	148	67.9	830	8	Adq97134 Human can
13	148	67.9	834	5	Aau77715 Diseased
14	148	67.9	854	6	Abp58332 Human cel
15	148	67.9	854	7	Adb75453 Prostate
16	148	67.9	911	7	Adb75455 Prostate
17	148	67.9	923	4	Abg01080 Novel hum
18	148	67.9	940	8	Adf66480 Human pro
19	148	67.9	940	8	Adf66138 Human pro
20	148	67.9	941	8	Adq97130 Human can
21	148	67.9	948	8	Ado55161 Protein #
22	148	67.9	954	8	Adq97125 Mouse can
23	148	67.9	955	8	Adq07990 Human Ned
24	148	67.9	958	8	Adq97128 Human can
25	148	67.9	975	2	Aaw93167 Human ZGG

26	148	67.9	995	6	AAE32719	Aae32719 KIAA0439
27	148	67.9	995	7	ADB75451	Prostate
28	148	67.9	995	7	AD660011	Human Pro
29	148	67.9	995	7	ADD45762	Human Pro
30	148	67.9	995	8	ADR04136	Human Ned
31	148	67.9	1000	8	ADQ97132	Human can
32	148	67.9	1071	4	ABG11772	Novel hum
33	142	65.1	37	2	AAW37638	Peptide c
34	142	65.1	38	7	ADB49230	Novel hum
35	138	63.3	38	2	AAR97681	Mouse Ned
36	138	63.3	38	2	AAW37681	Mouse Ned
37	138	63.3	1082	4	ABD49214	Mouse WW
38	133	61.0	38	7	ABE63355	Drosophila
39	132	60.6	38	2	ADB49215	Human WW
40	132	60.6	38	3	AAW37683	Yeast Rsp
41	132	60.6	38	4	AAW37683	RSPS/9C p
42	132	60.6	38	4	AAW37683	RSPS/9C p
43	128.5	58.9	37	5	AAU87969	Yeast WW
44	128.5	58.9	870	4	AAE05495	Human WW
45	128.5	58.9	870	6	AAE32722	Nedd4-4-11

ALIGNMENTS

RESULT 1

AAB83018  
ID AAB83018 standard; peptide; 38 AA.

AC AAB83018;

DT 25-JUN-2001 (first entry)

DE Human Nedda4 HECT E3 ubiquitin ligase WW domain #1.

Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
HECT E3 ubiquitin ligase; antiinflammatory; cytosolic; osteoparosis;  
antimicrobial; neuroprotective; transforming growth factor beta;  
TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
renal failure; neurodegeneration; fibrosis; WW domain; Nedda4.  
Homo sapiens.

WO200116604-A1.

08-MAR-2001.

29-AUG-2000; 2000WO-US023729.

30-AUG-1999; 99US-00385918.

(SIGN-) SIGNAL PHARM INC.

Hoekstra MF, Xie W, Murray BW, Mercurio FM;

WPI; 2001-327913/34.

Screening for modulators of TGF-beta and/or bone morphogenic protein  
(BMP) mediated signaling useful for treating cancer and osteoporosis by  
evaluating the ability of agents to modulate Smad protein degradation.

Disclosure; Page 13; 75pp; English.

The present sequence is the WW domain of a HECT (homologous to E6  
carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad  
PY motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.  
The sequence is provided in a specification relating to a method for  
screening for agents that modulate transforming growth factor (TGF)-beta  
and/or bone morphogenic protein (BMP)-mediated signalling. The method  
involves evaluating the effect of an agent on binding of HECT E3  
ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein

CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
 CC bone formation in a patient or treating a condition associated with  
 CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
 CC inhibit BMP-mediated signalling are useful for treating inflammation,  
 CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
 CC signalling are useful for stimulating bone anabolism as well as treating  
 CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
 CC that inhibit TGF-mediated signalling are useful for treating cancer,  
 CC inflammation, neurodegeneration and fibrosis

XX SQ Sequence 38 AA;

Query Match 100.0%; Score 218; DB 4; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
 Db 1 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38

RESULT 2

AAE32726

ID AAE32726 standard; protein; 733 AA.

XX AC AAE32726;

XX DT 24-MAR-2003 (first entry)

XX DE Human NEDD4 short form protein.

XX KW Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;  
 KW poliomyelitis; HIV; measles; protein therapy; human; NEDD4.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 260 /note= "Encoded by CAA"

FT Misc-difference 279 /note= "Encoded by AGC"

XX WO200290549-A2.

XX 14-NOV-2002.

XX PF 12-MAR-2002; 2002WO-IB002106.

XX PR 12-MAR-2001; 2001US-0275224P.

XX PR 31-JUL-2001; 2001US-0308958P.

XX PR 07-DEC-2001; 2001US-0340170P.

XX (PROT-) PROTEOLOGICS LTD.

XX Greener T, Moskowitz H, Reiss Y, Alroy I;  
 WPI; 2003-111976/10.  
 N-PSDB; AAD50460.

XX New protein complex comprising HECT-RCCL, viral maturation scaffolding  
 PT protein (VMSF), and/or HIV gag protein, useful for treating viral  
 PT infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,  
 PT or Ebola.

XX Disclosure; Fig 16; 150pp; English.

XX The invention relates to a method for modulation of viral maturation. The  
 CC invention also provides an isolated protein complex comprising a HECT-  
 CC RCCL selected from HECT-WW, HECT-RCCL, Gag protein, Gag late domain, P13,  
 CC actin, myosin, Hsp60, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM,  
 CC Grpase, E2 enzyme, tscl01, cullin, HECT1, HECT2, Nedd4 -like  
 CC protein or clathrin. The complexes, proteins, antibodies and methods are  
 CC useful for treating viral infections, such as lymphosarcoma, human

CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola  
 CC and for inhibiting budding in a subject. They are also useful in  
 CC diagnostic assays for determining whether a cell is infected with a virus  
 CC and for characterising the nature, progression and/or infectivity of the  
 CC infection. The invention is also useful in protein therapy. The present  
 CC sequence is human NEDD4 protein used to illustrate the method of the  
 CC invention

XX SQ Sequence 733 AA;

Query Match 100.0%; Score 218; DB 6; Length 733;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-20;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
 Db 191 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 228

RESULT 3

ADQ07991

ID ADQ07991 standard; protein; 900 AA.

XX AC ADQ07991;

XX DT 07-OCT-2004 (first entry)

XX DE Human NEDD4 protein.

XX KW MHCAT; modifier of beta-catenin; drug screening;  
 KW beta-catenin pathway modulator; wingless signalling pathway;  
 KW Wg signalling pathway; Wnt signalling pathway; cell cycle progression;  
 KW diagnosis; cancer; angiogenic disorder; apoptotic disorder;  
 KW cell proliferation disorder; cytostatic; gene therapy; human; NEDD4;  
 KW neural precursor cell expressed developmentally downregulated 4.

XX OS Homo sapiens.

XX WO2004061123-A2.

XX PD 22-JUL-2004.

XX PF 29-DEC-2003; 2003WO-US041404.

XX PR 30-DEC-2002; 2002US-0436974P.

XX PA (EXEL-) EXELIXIS INC.

XX Francis-Lang H, Winter CG, Ventura RBA, Lickteig K;  
 WPI; 2004-534389/51.  
 N-PSDB; ADQ07980.  
 REFSEQ; XM\_046129.4.

XX Identifying a candidate beta-catenin MHCAT pathway modulating agent,  
 PT useful for diagnosing and treating cancer, comprises providing an assay  
 PT system comprising a MHCAT polypeptide or nucleic acid.

XX Example II; SEQ ID NO 22; 113pp; English.

XX The invention relates to the use of MHCAT (modifier of beta-catenin)  
 CC polypeptides or nucleic acids for identifying a candidate beta-catenin  
 CC pathway modulating agent or for modulating a beta-catenin pathway of a  
 CC cell. The invention also relates to a method of modulating a beta-catenin  
 CC pathway in a cell; and diagnosing a disease in a patient by detecting  
 CC MHCAT expression. The invention is based on the finding that MHCAT  
 CC orthologues are able to modify the beta-catenin pathway in Drosophila  
 CC cells. Beta-catenin (also known as Armadillo in Drosophila) plays a role  
 CC in cell signalling via the Wingless (Wg)/Wnt signalling pathway which is  
 CC involved in cell cycle progression. Beta-catenin levels are tightly  
 CC regulated by a complex containing the tumour suppressor proteins APC  
 CC (adenomatous polyposis coli) and Axin and the serine/threonine kinase  
 CC GSK3-beta. In the absence of these proteins, beta-catenin accumulates in



the nucleus where it acts as a transcriptional co-activator with TCF for the induction of target genes such as the cell-cycle regulators cyclin D1 and c-Myc. The Wingless/beta-catenin signalling pathway is frequently mutated in human cancers, particularly those of the colon. Beta-catenin pathway modulating agents identified using the methods of the invention are useful as therapeutic agents for treating disorders associated with defective or impaired beta-catenin function and/or MCAR function such as cancer, or angiogenic, apoptotic or cell proliferation disorders. Proteins that interact with MCAR are also useful in detecting and providing information about the function of MCAR proteins. The present sequence represents a human NECD4 (neural precursor cell expressed, developmentally downregulated 4) protein, identified as an MCAR polypeptide in an example of the invention.

XX SQ Sequence 900 AA;

Query Match 100.0%; Score 218; DB 8; Length 900;  
Best Local Similarity 100.0%; Pred. No. 8.2e-20;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
|||||  
Db 191 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 228  
|||||

RESULT 4

AAAY25170  
ID AAY25170 standard; protein; 927 AA.

XX AC AAY25170;

XX DT 09-SEP-1999 (first entry)

XX DE Human KIAAN ligase protein fragment.

XX KW Inhibitor; ubiquitin-mediated proteolysis; IkappaB polypeptide; HECT; ubiquitin-conjugating system; Homologous to E6-Ap carboxyl terminus; ligase; ubiquitin; ubiquitination; RSC ligase; KIAAN ligase; UBC4; MAD3; modulator; treatment; proliferative disorder; apoptosis; sepsis; differentiative disorder; viral infection; tissue wasting disorder; cachexia; malignancy; inflammatory disease; parasitic disease; tuberculosis; IL-2 therapy; rheumatoid arthritis; rheumatoid spondylitis; osteoarthritis; gouty arthritis; respiratory distress syndrome; cerebral malaria; chronic pulmonary inflammatory disease; silicosis; pulmonary sarcoidosis; bone resorption disease; reperfusion injury; graft versus host reaction; allograft rejection; Crohn's disease; ulcerative colitis; pyresis; multiple sclerosis; autoimmune diabetes; systemic lupus erythematosus; leprosy; AIDS.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT Protein 1..927  
XX FT /note= "partial sequence, no start codon given"

XX PN WO9904033-A1.

XX PD 28-JAN-1999.

XX PF 16-JUL-1998; 98WO-US014638.

XX PR 16-JUL-1997; 97US-00895601.

XX PA (MITO-) MITOTIX INC.

XX PI Beer-Romero P, Strack PR, Glass SJ, Rolfe M;

XX DR WPI; 1999-132274/11.

XX DR N-PSDB; AAX78494.

XX PT Identifying modulators of I-kappa-B proteolysis - used to develop products for treating e.g. proliferative and/or differentiative disorders, infections, tissue wasting, cachexia or AIDS.

XX Disclosure; Page 68-71; 79pp; English.

XX This invention describes a novel assay identifying an inhibitor of ubiquitin-mediated proteolysis of a IkappaB polypeptide. The method comprises: (a) providing a ubiquitin-conjugating system including the IkappaB polypeptide and a HECT (Homologous to E6-Ap Carboxyl Terminus) ligase and ubiquitin, to promote ubiquitination of the IkappaB polypeptide by the HECT ligase; (b) contacting the ubiquitin-conjugating system with a candidate agent; (c) measuring a level of ubiquitination of the polypeptide in the presence of the candidate agent and (d) comparing the measured level of ubiquitination in the presence of the candidate agent with ubiquitination of the IkappaB polypeptide in the absence of the candidate agent; where a statistically significant decrease in ubiquitination of the IkappaB polypeptide in the presence of the candidate agent is indicative of an inhibitor of ubiquitination of the IkappaB polypeptide. The assays can be used to identify compounds which modulate binding and/or ubiquitination of an IkappaB (or other cellular or viral substrate) by a HECT ligase, such as RSC or KIAAN. Such modulators can be used e.g. in the treatment of proliferative and/or differentiative disorders, to modulate apoptosis, in the treatment of viral infections, and in the treatment of tissue wasting disorders e.g. cachexia secondary to infection or malignancy, cachexia secondary to human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis and other arthritic conditions, sepsis, respiratory distress syndrome, cerebral malaria, chronic pulmonary inflammatory disease, silicosis, pulmonary sarcoidosis, bone resorption diseases, reperfusion injury, graft versus host reaction, allograft rejections, Crohn's disease, ulcerative colitis, or pyresis, in addition to a number of autoimmune diseases such as multiple sclerosis, autoimmune diabetes, systemic lupus erythematosus; and ENL in leprosy, HIV, and AIDS. This sequence represents a human KIAAN ligase which is used in the method of the invention

XX SQ Sequence 927 AA;

Query Match 100.0%; Score 218; DB 2; Length 927;  
Best Local Similarity 100.0%; Pred. No. 8.5e-20;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
|||||  
Db 218 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 255  
|||||

RESULT 5

AAE32725  
ID AAE32725 standard; protein; 927 AA.

XX AC AAE32725;

XX DT 24-MAR-2003 (first entry)

XX DE Human NECD4 long form protein.

XX KW Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis; poliomyelitis; HIV; measles; protein therapy; human; NECD4.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT Domain 381..406  
XX FT /note= "WW domain"

XX FT Domain 821..923  
XX FT /note= "HECT domain"

XX PN WO200290549-A2.

XX PD 14-NOV-2002.

XX PF 12-MAR-2002; 2002WO-IB002106.

XX

PR 12-MAR-2001; 2001US-0275224P.  
PR 31-JUL-2001; 2001US-0308958P.  
PR 07-DEC-2001; 2001US-0340170P.  
XX (PROT-) PROTEOLOGICS LTD.  
PA Greener T, Moskowitz H, Reiss Y, Alroy I;  
PI WPI; 2003-111976/10.  
XX N-PSDB; AAD50459.  
DR  
DR  
XX New protein complex comprising HECT-RCCL1, viral maturation scaffolding protein (VMSF), and/or HIV gag protein, useful for treating viral infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles, or Ebola.  
XX Disclosure; Fig 15; 150pp; English.  
XX The invention relates to a method for modulation of viral maturation. The invention also provides an isolated protein complex comprising a HECT-RCCL1 selected from HECT-WW, HECT-RCCL1, Gag protein, Gag late domain, P13, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, Grpase, E2 enzyme, tsg101, cullin, HEC1, HEC2, HEC3, Nedd4-like protein or clathrin. The complexes, proteins, antibodies and methods are useful for treating viral infections, such as lymphosarcoma, human immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola and for inhibiting budding in a subject. They are also useful in diagnostic assays for determining whether a cell is infected with a virus and for characterizing the nature, progression and/or infectivity of the infection. The invention is also useful in protein therapy. The present sequence is human NEDD4 protein used to illustrate the method of the invention  
XX  
XX Sequence 927 AA;  
Query Match 100.0%; Score 218; DB 6; Length 927;  
Best Local Similarity 100.0%; Pred. No. 8.5e-20;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
Db 218 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 255  
RESULT 6  
AAB83019  
ID AAB83019 standard; peptide; 38 AA.  
XX  
AC AAB83019;  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Human Nedd4 HECT E3 ubiquitin ligase WW domain #2.  
XX  
KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; WW domain; Nedd4.  
XX  
OS Homo sapiens.  
XX  
XX W0200116604-A1.  
XX  
XX 08-MAR-2001.  
XX  
XX 29-AUG-2000; 2000WO-US023729.  
XX  
XX 30-AUG-1999; 99US-00385918.  
XX  
XX (SIGN-) SIGNAL PHARM INC.  
XX

PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
XX WPI; 2001-327913/34.  
XX Screening for modulators of TGF-beta and/or bone morphogenic protein (BMP) mediated signaling useful for treating cancer and osteoporosis by evaluating the ability of agents to modulate Smad protein degradation.  
XX Disclosure; Page 13; 75pp; English.  
XX The present sequence is the WW domain of a HECT (homologous to E6 carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad PY motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.  
CC The sequence is provided in a specification relating to a method for screening for agents that modulate transforming growth factor (TGF)-beta and/or bone morphogenic protein (BMP)-mediated signalling. The method involves evaluating the effect of an agent on binding of HECT E3 ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein HECT E3 ubiquitin ligase activity. The method is useful for stimulating bone formation in a patient or treating a condition associated with insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that inhibit BMP-mediated signalling are useful for treating inflammation, ageing, cancer and infectious diseases. Agents that augment BMP-mediated signalling are useful for stimulating bone anabolism as well as treating broken bones, osteoporosis, and acute or chronic renal failure. Agents that inhibit TGF-mediated signalling are useful for treating cancer, inflammation, neurodegeneration and fibrosis  
XX  
XX Sequence 38 AA;  
Query Match 95.9%; Score 209; DB 4; Length 38;  
Best Local Similarity 97.4%; Pred. No. 3.6e-20;  
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SPLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
Db 1 SGLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
RESULT 7  
ADB49212  
ID ADB49212 standard; peptide; 38 AA.  
XX  
AC ADB49212;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Mouse WW binding protein Ned4 WW domain #1.  
XX  
KW WW domain; drug candidate screening; drug discovery; drug modification;  
KW drug refinement; immunogen; WW binding protein; mouse; ned4, WW domain.  
XX  
OS Mus sp.  
XX  
XX US2003077577-A1.  
XX  
XX 24-APR-2003.  
XX  
XX 28-JUN-2002; 2002US-00185050.  
XX  
XX 03-APR-1996; 96US-00630916.  
PR 03-APR-1997; 97US-00826516.  
XX  
XX (PIRO/) PIROZZI G.  
PA (KAYB/) KAY B K.  
PA (FOWL/) FOWLKES D M.  
XX  
XX Pirozzi G, Kay BK, Fowlkes DM;  
XX  
XX WPI; 2003-635075/60.  
XX  
XX Novel purified polypeptide comprising WW domain, useful for drug

PT discovery, modification and refinement, for discovering polypeptides  
 PT involved in pharmacological activities, or as an immunogen to generate  
 XX antibodies.  
 XX Example; Fig 5; 133pp; English.  
 XX The invention describes a purified polypeptide (I) comprising a WW domain  
 CC which has a sequence (S1) selected from 11 sequences fully defined in the  
 CC specification, a sequence (S2) selected from 48 sequences fully defined  
 CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725  
 CC amino acids fully defined in the specification. (I) is useful for  
 CC screening a potential drug candidate, by allowing (I) to come into  
 CC contact with at least one recognition unit having a selective affinity  
 CC for the WW domain in (I), in the presence of an amount of a potential  
 CC drug candidate, such that (I) and the recognition unit are capable of  
 CC interacting when brought into contact with one another in the absence of  
 CC the drug candidate, and determining the effect, if any, of the presence  
 CC of the amount of the drug candidate on the interaction of (I) with the  
 CC recognition unit. (I) is useful for drug discovery, modification and  
 CC refinement, for discovering polypeptides involved in pharmacological  
 CC activities, or as an immunogen to generate antibodies. This is the amino  
 CC acid sequence of WW binding protein Ned4 WW domain.  
 XX  
 SQ Sequence 38 AA;  
 Query Match 95.9%; Score 209; DB 7; Length 38;  
 Best Local Similarity 97.4%; Pred. No. 3.6e-20;  
 Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
 Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
 RESULT 8  
 AAR97680  
 ID AAR97680 standard; protein; 38 AA.  
 XX  
 AC AAR97680;  
 XX  
 DT 30-AUG-1996 (first entry)  
 DE Mouse Nedda4 WW domain-1.  
 XX  
 XX WW domain; signal transduction; diagnosis; gene therapy;  
 KW Yes proto-oncogene associated protein; YAP; Nedda4.  
 KW  
 XX Mus sp.  
 XX  
 XX WO9617061-A1.  
 XX  
 PD 06-JUN-1996.  
 XX  
 XX 30-NOV-1995; 95WO-US015512.  
 XX  
 XX 01-DEC-1994; 94US-00348518.  
 PR  
 PR 07-JUN-1995; 95US-00476509.  
 XX  
 XX (UVRQ ) UNIV ROCKEFELLER.  
 PA (DELB-) DELBRUECK CENT MOLECULAR MEDICINE MAX.  
 XX  
 XX Sudol M, Chen H, Bork P;  
 XX  
 XX WPI; 1996-286829/29.  
 XX  
 XX DNA encoding Yes proto-oncogene associated protein - used to modulate  
 PT intracellular signal transduction e.g. for treatment of muscular  
 PT dystrophy.  
 XX  
 XX Claim 10; Fig 12; 126pp; English.  
 PS  
 XX WW domains (AAR97673-92) were identified in a number of proteins involved  
 CC in signalling or regulatory functions. The WW domain was initially

CC identified in the Yes proto-oncogene associated proteins (YAPs) of  
 CC chicken, human and mouse (see also AAR97663-70 and AAR97672). A consensus  
 CC sequence is given in AAR97671. Mouse Nedda4 plays a role in embryonic  
 CC development and CNS differentiation. The Nedda4 WW domains (AAR97680-82)  
 CC can be expressed in bacterial, yeast, insect or mammalian cells, and used  
 CC to identify WW domain ligands. They can be introduced into cells, either  
 CC directly or by gene therapy, to increase the level of signal transduction  
 XX  
 SQ Sequence 38 AA;  
 Query Match 93.1%; Score 203; DB 2; Length 38;  
 Best Local Similarity 89.5%; Pred. No. 2.3e-19;  
 Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
 Db 1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPSPDDDL 38  
 RESULT 9  
 AAB21978  
 ID AAB21978 standard; peptide; 38 AA.  
 XX  
 AC AAB21978;  
 XX  
 DT 02-JAN-2001 (first entry)  
 DE Nedda4/mouse peptide containing a WW-domain #2.  
 XX  
 KW WW-domain; protein-protein interaction; cell growth regulation;  
 KW protein degradation regulation; Alzheimer's; Dementia pugilistica;  
 KW Down's syndrome; Parkinson's disease; Pick's; neurodegenerative;  
 KW microtubule assembly; tau; hyperplasia; neoplasia; malignancy; psoriasis;  
 KW retinosis; atherosclerosis; leukaemia; lymphoma; papiloma;  
 KW pulmonary fibrosis; rheumatoid arthritis; multiple sclerosis;  
 KW muscular dystrophy; mouse.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200048621-A2.  
 PN  
 PD 24-AUG-2000.  
 XX  
 PF 18-FEB-2000; 2000WO-US004278.  
 XX  
 PR 18-FEB-1999; 99US-00252404.  
 XX  
 XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 PA  
 XX Lu KP, Zhou XZ;  
 XX  
 XX WPI; 2000-594014/56.  
 DR  
 XX  
 PT Mediating protein-protein interactions, useful for regulating cell growth  
 PT and for treating neurodegenerative disorders, comprises modulating  
 PT binding of WW domain containing polypeptide with phosphorylated ligand.  
 XX  
 XX Disclosure; Fig 6; 82pp; English.  
 XX  
 XX The present invention relates to a method for mediating protein-protein  
 CC interaction, which comprises modulating the binding of a WW-domain  
 CC containing peptide with a phosphorylated ligand e.g. tau. WW-domains are  
 CC highly conserved regions of approximately 40 amino acid residues with two  
 CC invariant tryptophans (W) in a triple stranded beta-sheet. The present  
 CC sequence is one such WW-domain. When a WW-domain containing peptide is  
 CC phosphorylated at serine or threonine residues, dephosphorylation of  
 CC ligands bound to the peptide is inhibited. The present peptide may be  
 CC useful for mediating protein-protein interaction, regulating cell growth,  
 CC regulating protein degradation, restoring the function of tau to bind  
 CC microtubules and promote or restore microtubule assembly in  
 CC neurodegenerative diseases e.g. Alzheimer's, Dementia pugilistica, Down's  
 CC syndrome, Parkinson's disease, Pick's disease, multiple sclerosis,  
 CC muscular dystrophy, Corticobasal degeneration, Frontotemporal dementias,



```
RESULT 12
ADQ97134
ID ADQ97134 standard; protein; 830 AA.
XX AC
XX ADQ97134;
XX DT
XX 07-OCT-2004 (first entry)
XX DE
XX Human cancer associated sequence HP4-08-005, SEQ ID 110.
XX KW
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.
XX OS
XX Homo sapiens.
XX PN
XX WO2004060304-A2.
XX PD
XX 22-JUL-2004.
XX PF
XX 22-DEC-2003; 2003WO-US041389.
XX PR
XX 27-DEC-2002; 2002US-00330773.
XX PA
XX (SAGR-) SAGRES DISCOVERY INC.
XX PI
XX Morris DW, Malandro MS;
XX DR
XX WPI; 2004-543781/52.
XX PT
XX New isolated cancer associated nucleic acids comprising at least 10
XX contigous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX FS
XX Claim 1; SEQ ID NO 110; 199pp; English.
XX CC
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX Sequence 830 AA;

Query Match 67.9%; Score 148; DB 8; Length 830;
Best Local Similarity 71.4%; Pred. No. 1.4e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PLPPGWEERQDILGRTYYVNHESRRRTQWKRTPOD 36
|||||||:|||||||:|||||:|
Db 153 PLPPGWEKVDNLGRTYYVNHNNRTTQWHRPSLMD 187

RESULT 13
AAU77715
ID AAU77715 standard; protein; 834 AA.
XX AC
XX AAU77715;
XX DT
XX 05-JUN-2002 (first entry)
XX DE
XX Diseased kidney tissue associated novel human protein #2.
XX KW
XX Kidney disease; nephrotropic; renal disease; Iga nephropathy; Igan;
XX asymptomatic microscopic haematuria; proteinuria;
XX episodic gross haematuria; necrotising crescentic glomerulonephritis;
XX NCGN; minimal change disease; Min Ch; renal hypertrophy; diabetes;
XX end-stage renal disease; ESRD; kidney function.
XX OS
XX Homo sapiens.
XX PN
XX WO200212439-A2.
XX PD
XX 14-FEB-2002.

us-09-385-918-3.rag
03-AUG-2001; 2001WO-US024635.
03-AUG-2000; 2000US-0222731P.
(GENE-) GENE LOGIC INC.
Sun H, Munger WE, Falk RJ;
WPI; 2002-257475/30.
N-PSDB; ABK11668.
Novel human genes that are differentially expressed in renal biopsy
samples from patients with kidney diseases useful for diagnosing renal
diseases such as Iga nephropathy or necrotizing crescentic
glomerulonephritis.
Claim 16; Page 62-65; 67pp; English.
The invention describes an isolated nucleic acid molecule (I) that is
expressed in diseased kidney tissue compared to normal kidney tissue. (I)
and its encoded protein (II) are useful for diagnosing a renal disease
state e.g. Iga nephropathy (IGAN) which presents as asymptomatic
microscopic haematuria and/or proteinuria or episodic gross haematuria,
necrotising crescentic glomerulonephritis (NCGN), minimal change disease
(Min Ch) or renal hypertrophy (the cause of end-stage renal disease and
early manifestation of diabetes), in a subject. (II) is useful: as a
target for identifying agents which modulate the expression or activity
of (I); for identifying binding partners; as novel target for screening
synthetic small molecules and combinatorial or naturally occurring
compound libraries to discover novel therapeutics to regulate kidney
function; as an antigen to raise polyclonal or monoclonal antibodies, and
as a therapeutic agent or target. Agents that modulate, up-or-down-
regulate the expression of (II) or agents that agonise or antagonise the
activity of the protein are useful to modulate biological and pathologic
purposes associated with proteins function and activity, for e.g. kidney
damage or end-stage renal disease (ESRD). This is the amino acid sequence
of the diseased kidney associated novel human protein #2, described in
the invention
XX SQ
XX Sequence 834 AA;

Query Match 67.9%; Score 148; DB 5; Length 834;
Best Local Similarity 71.4%; Pred. No. 1.4e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PLPPGWEERQDILGRTYYVNHESRRRTQWKRTPOD 36
|||||||:|||||||:|||||:|
Db 73 PLPPGWEKVDNLGRTYYVNHNNRTTQWHRPSLMD 107

RESULT 14
ABP58332
ID ABP58332 standard; protein; 854 AA.
XX AC
XX ABP58332;
XX DT
XX 07-APR-2003 (first entry)
XX DE
XX Human cell growth, differentiation and death protein CGDD-3.
XX KW
XX CGDD-3; cell growth; cell differentiation; cell death; human; cytostatic;
XX antiarteriosclerotic; hepatotropic; antinflammatory; antiporotic;
XX antinaemic; ophthalmological; auditory; anticonvulsant;
XX cerebroprotective; nootropic; neuroprotective; antiparkinsonian;
XX neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
XX antiasthmatic; antihyroid; antidiabetic; dermatological; nephrotropic;
XX antineumatic; antiarthritic; anticancer; vulnerary; virucide;
XX antibacterial; fungicide; antiparasitic; protozoacide; antihelminthic;
XX antinfertility; gynaecological; ubiquitin protein ligase; enzyme;
XX gene therapy.
XX OS
XX Homo sapiens.
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 81.011 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-3  
 Perfect score: 218  
 Sequence: 1 SPLPPGWEERODILGRTVVYNHESRRTOWKRPPODNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10D\_PUBCOMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US10E\_PUBCOMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pcp.\*
- 19: /cgn2\_6/ptodata/1/pubaa/US11A\_PUBCOMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/pubaa/US11\_NEW\_PUB.pcp.\*
- 21: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pcp.\*
- 22: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score		Query %		Length	DB	ID	Description
	Score	Match						
1	218	100.0			38	14	US-10-307-956-3	Sequence 3, Appli
2	218	100.0			733	14	US-10-097-534-16	Sequence 16, Appl
3	218	100.0			927	14	US-10-097-534-15	Sequence 15, Appl
4	209	95.9			38	14	US-10-185-050-18	Sequence 18, Appl
5	209	95.9			38	14	US-10-307-956-4	Sequence 4, Appli
6	203	93.1			38	14	US-10-256-865-36	Sequence 36, Appl
7	203	93.1			38	14	US-10-185-050-17	Sequence 17, Appl
8	203	93.1			38	16	US-10-687-361-11	Sequence 11, Appl
9	203	93.1			38	17	US-10-716-379-11	Sequence 11, Appl
10	203	93.1			38	17	US-10-616-410-11	Sequence 11, Appl
11	203	93.1			38	17	US-10-648-631-11	Sequence 11, Appl

12	148	67.9	834	14	US-10-313-955-6	Sequence 6, Appl
13	148	67.9	854	14	US-10-205-823-277	Sequence 277, App
14	148	67.9	854	14	US-10-287-218-3	Sequence 3, Appl
15	148	67.9	854	16	US-10-474-291-3	Sequence 3, Appl
16	148	67.9	854	20	US-11-051-454-277	Sequence 277, App
17	148	67.9	911	14	US-10-205-823-279	Sequence 279, App
18	148	67.9	911	20	US-11-051-454-279	Sequence 279, App
19	148	67.9	923	18	US-10-450-763-31439	Sequence 31439, A
20	148	67.9	975	16	US-10-779-271-3	Sequence 3, Appl
21	148	67.9	995	14	US-10-097-534-9	Sequence 9, Appl
22	148	67.9	995	14	US-10-205-823-275	Sequence 275, App
23	148	67.9	995	20	US-11-051-454-275	Sequence 275, App
24	148	67.9	1071	18	US-10-450-763-42131	Sequence 42131, A
25	142	65.1	38	14	US-10-485-050-36	Sequence 36, Appl
26	138	63.3	38	14	US-10-185-050-20	Sequence 20, Appl
27	138	63.3	1082	20	US-11-097-143-16857	Sequence 16857, A
28	133	61.0	38	14	US-10-185-050-21	Sequence 21, Appl
29	132	60.6	38	14	US-10-256-865-37	Sequence 37, Appl
30	132	60.6	38	14	US-10-485-050-19	Sequence 19, Appl
31	132	60.6	38	14	US-10-307-956-7	Sequence 7, Appl
32	132	60.6	38	16	US-10-687-361-12	Sequence 12, Appl
33	132	60.6	38	17	US-10-716-379-12	Sequence 12, Appl
34	132	60.6	38	17	US-10-616-410-12	Sequence 12, Appl
35	132	60.6	38	17	US-10-648-631-12	Sequence 12, Appl
36	128.5	58.9	37	17	US-10-785-819-165	Sequence 165, App
37	128.5	58.9	870	14	US-10-097-534-12	Sequence 12, Appl
38	128.5	58.9	870	16	US-10-723-860-2167	Sequence 2167, Ap
39	128.5	58.9	906	14	US-10-185-050-48	Sequence 48, Appl
40	128	58.7	38	14	US-10-185-050-35	Sequence 35, Appl
41	128	58.7	38	17	US-10-785-819-164	Sequence 164, App
42	128	58.7	474	10	US-09-774-639-371	Sequence 371, App
43	128	58.7	474	10	US-09-969-730-249	Sequence 249, App
44	128	58.7	474	15	US-10-621-363-249	Sequence 249, App
45	128	58.7	739	14	US-10-097-534-10	Sequence 10, Appl

## ALIGNMENTS

```

RESULT 1
US-10-307-956-3
; Sequence 3, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-3

```

```

Query Match      100.0%; Score 218; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. NO. 9.le-21;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 SP1PPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38
Db  1 SP1PPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

```

## RESULT 2

US-10-097-534-16  
; Sequence 16, Application US/10097534  
; Publication No. US20030049607A1  
; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; TITLE OF INVENTION: MATURATION  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/340,170  
; PRIOR FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 733  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-534-16

Query Match 100.0%; Score 218; DB 14; Length 733;  
Best Local Similarity 100.0%; Pred. No. 1.9e-19;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38  
|||||  
Db 191 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 228

RESULT 3  
US-10-097-534-15  
; Sequence 15, Application US/10097534  
; Publication No. US20030049607A1  
; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; TITLE OF INVENTION: MATURATION  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/340,170  
; PRIOR FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 927  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-534-15

Query Match 100.0%; Score 218; DB 14; Length 927;  
Best Local Similarity 100.0%; Pred. No. 2.4e-19;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38  
|||||  
Db 218 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 255

RESULT 4  
US-10-185-050-18  
; Sequence 18, Application US/10185050  
; Publication No. US20030077577A1  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; Kay, Brian K.  
; Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 233  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/185,050  
; FILING DATE: 28-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/826,516  
; FILING DATE: 03-Apr-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-208-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-10-185-050-18  
  
Query Match 95.9%; Score 209; DB 14; Length 38;  
Best Local Similarity 97.4%; Pred. No. 1.3e-19;  
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38  
|||||  
Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38  
  
RESULT 5  
US-10-307-956-4  
; Sequence 4, Application US/10307956  
; Publication No. US20030119072A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoeckstra, Merl F.  
; APPLICANT: Xie, Weilin  
; APPLICANT: Murray, Brion  
; APPLICANT: Mercurio, Frank  
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL  
; TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS  
; FILE REFERENCE: 860098.433  
; CURRENT APPLICATION NUMBER: US/10/307,956  
; CURRENT FILING DATE: 2002-12-02  
; PRIOR APPLICATION NUMBER: US/09/385,918  
; PRIOR FILING DATE: 1999-08-30



```
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-4

Query Match      95.9%; Score 209; DB 14; Length 38;
Best Local Similarity 97.4%; Pred. No. 1.3e-19;
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPGWEERQDILGRYYVNHESRRTQWKRTPODNL 38
Db 1 SGLPPGWEERQDILGRYYVNHESRRTQWKRTPODNL 38

RESULT 6
US-10-256-865-36
; Sequence 36, Application US/10256865
; Publication No. US20030049672A1
; GENERAL INFORMATION:
; APPLICANT: Xiao Zhen Zhou
; TITLE OF INVENTION: Methods and Compositions for Regulating
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 1440.1034-000
; CURRENT APPLICATION NUMBER: US/10/256,865
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: PRIOR APPLICATION NUMBER: US/09/252,404
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 38
; TYPE: PRT
; ORGANISM: mouse
US-10-256-865-36

Query Match      93.1%; Score 203; DB 14; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPGWEERQDILGRYYVNHESRRTQWKRTPODNL 38
Db 1 SPLPGWEERQDVLGRYYVNHESRRTQWKRPSPDDL 38

RESULT 7
US-10-185-050-17
; Sequence 17, Application US/10185050
; Publication No. US2003007577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
```

```
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-185-050-17

Query Match      93.1%; Score 203; DB 14; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPGWEERQDILGRYYVNHESRRTQWKRTPODNL 38
Db 1 SPLPGWEERQDVLGRYYVNHESRRTQWKRPSPDDL 38

RESULT 8
US-10-687-361-11
; Sequence 11, Application US/10687361
; Publication No. US20040101896A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Lu, Kun Ping
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-078
; CURRENT APPLICATION NUMBER: US/10/687,361
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-687-361-11

Query Match      93.1%; Score 203; DB 16; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPGWEERQDILGRYYVNHESRRTQWKRTPODNL 38
Db 1 SPLPGWEERQDVLGRYYVNHESRRTQWKRPSPDDL 38

RESULT 9
US-10-716-379-11
; Sequence 11, Application US/10716379
; Publication No. US20050027107A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Kun Ping, Lu
; APPLICANT: Hanes, Steven D.
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-085
; CURRENT APPLICATION NUMBER: US/10/716,379
```

```

; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-716-379-11

Query Match          93.1%; Score 203; DB 17; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPUPPGWEERQDILGRYYVYNNHESRRQTQWKRPDPQDL 38
      |||||||||:|||||:|||||:|||||:|||||:|
DB 1 SPUPPGWEERQDVLGRYYVYNNHESRRQTQWKRPSPDDDL 38

RESULT 10
US-10-616-410-11
; Sequence 11, Application US/10616410
; Publication No. US20050033032A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Kun Ping, Lu
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-043
; CURRENT APPLICATION NUMBER: US/10/616,410
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US 09/275,900
; PRIOR FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-616-410-11

Query Match          93.1%; Score 203; DB 17; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPUPPGWEERQDILGRYYVYNNHESRRQTQWKRPDPQDL 38
      |||||||||:|||||:|||||:|||||:|||||:|
DB 1 SPUPPGWEERQDVLGRYYVYNNHESRRQTQWKRPSPDDDL 38

RESULT 11
US-10-648-631-11
; Sequence 11, Application US/10648631
; Publication No. US20050049404A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Kun Ping, Lu
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-044
; CURRENT APPLICATION NUMBER: US/10/648,631
; CURRENT FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-648-631-11

Query Match          93.1%; Score 203; DB 17; Length 38;
Best Local Similarity 89.5%; Pred. No. 8e-19;
Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-277

```

```

Query Match      67.9%; Score 148; DB 14; Length 854;
Best Local Similarity 71.4%; Pred. No. 2.6e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      2 PLPPGWEERQDILGRYYVNHESRRQWKRPQD 36
        |||||:|||||:|||||:|||||:|||||:
Db      73 PLPPGWEKVDNLGRYYVNHNNRTTQWHRPSLMD 107

```

```

RESULT 14
US-10-287-218-3
; Sequence 3, Application US/10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BUFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727

```

```

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 3660046CD1
US-10-287-218-3

```

```

Query Match      67.9%; Score 148; DB 14; Length 854;
Best Local Similarity 71.4%; Pred. No. 2.6e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      2 PLPPGWEERQDILGRYYVNHESRRQWKRPQD 36
        |||||:|||||:|||||:|||||:|||||:
Db      73 PLPPGWEKVDNLGRYYVNHNNRTTQWHRPSLMD 107

```

```

RESULT 15
US-10-474-291-3
; Sequence 3, Application US/10474291
; Publication No. US20040132043A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BUFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Javalaxmi
; APPLICANT: REDDY, Roopa M.; SANJANWALA, Madhusudan M.
; APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USN
; CURRENT APPLICATION NUMBER: US/10/474,291
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27

```

```

; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 3660046CD1
US-10-474-291-3

```

Query Match 67.9%; Score 148; DB 16; Length 854;  
Best Local Similarity 71.4%; Pred. No. 2.6e-10;  
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PLPGWEERQDILGRYYVNHESRRRTQWKRTPD 36  
 |||||:|||||||:|||||  
 Db 73 PLPGWEEKVDNLGRYYVNNHNRRTQWHRPSLMD 107

Search completed: October 13, 2005, 14:50:00  
Job time : 82.011 secs



F;402-439/Domain: WW repeat homology <WW2>  
F;459-496/Domain: WW repeat homology <WW3>  
F;555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 91.3%; Score 199; DB 2; Length 887;  
Best Local Similarity 86.8%; Pred. No. 2.1e-17;  
Matches 33; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPLPGWEERQDILGRYYVNHESRRTQWKRPDQNL 38  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 246 SPLPGWEERQDVLGRYYVNHESRTTQWRKRPEDDL 283

## RESULT 3

T46412  
ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)  
N;Alternate names: hypothetical protein DKF2p434P2422.1  
C;Species: Homo sapiens (man)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T46412  
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23034  
A;Accession: T46412  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-820 <AA>  
A;Cross-references: UNIPROT:Q9NT88; EMBL:AL137469  
A;Experimental source: adult testis; clone DKF2p434P2422  
C;Genetics:  
A;Gene: GDB:NEDD4  
A;Note: DKF2p434P2422.1  
C;Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin-protein ligase  
C;Keywords: ligase  
F;142-179/Domain: WW repeat homology <WW1>  
F;342-379/Domain: WW repeat homology <WW2>  
F;393-430/Domain: WW repeat homology <WW3>  
F;489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 67.9%; Score 148; DB 2; Length 820;  
Best Local Similarity 71.4%; Pred. No. 6.4e-11;  
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PLPGWEERQDILGRYYVNHESRRTQWKRPDQD 36  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 143 PLPGWEKVDNLGRYYVNHNNRTQWHRPSLMD 177

## RESULT 4

S43217  
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)  
N;Alternate names: E8-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein  
C;Species: Saccharomyces cerevisiae  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S43217; S50628; S70050  
R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993  
A;Reference number: S30812  
A;Accession: S43217  
A;Molecule type: DNA  
A;Residues: 1-809 <MUL>  
A;Cross-references: UNIPROT:P39940; GB:U18916; EMBL:L11119; NID:g1384128; PIDN:AAC03223.  
R;Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A;Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda  
A;Reference number: S50628  
A;Accession: S50628  
A;Molecule type: DNA  
A;Residues: 1-809 <DIE>

A;Cross-references: EMBL:U18916; NID:g1384128; PIDN:AAC03223.1; PID:g603364; GSPDB:GN000  
R;Hein, C.; Springael, J.Y.; Volland, C.; Hagenauer-Tsapis, R.; Andre, B.  
Mol. Microbiol. 18, 77-87, 1995  
A;Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur4

A;Reference number: S70050; MUID:96154942; PMID:8596462  
A;Accession: S70050  
A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
A;Molecule type: DNA  
C;Genetics: 1-101 <HE1>  
A;Residues:  
A;Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125w  
A;Cross-references: MIPS:YER125w; SGD:S0000927  
A;Map position: SR  
A;Function:  
A;Description: involved in endocytosis of Gap1 protein and Fur4 protein; binds and ubiquitinates  
C;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C;Keywords: ligase  
F;229-266/Domain: WW repeat homology <WW1>  
F;331-368/Domain: WW repeat homology <WW2>  
F;387-424/Domain: WW repeat homology <WW3>

Query Match 60.6%; Score 132; DB 1; Length 809;  
Best Local Similarity 71.0%; Pred. No. 7e-09;  
Matches 22; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPFGWEERQDILGRYYVNHESRRTQWKRP 33  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 231 LPFGWEERTDNFGRTYYVDHNTTRTTWKRP 261

## RESULT 5

T39585  
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T39585  
R;Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1997  
A;Reference number: Z21865  
A;Accession: T39585  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-786 <VOL>  
A;Cross-references: UNIPROT:O14326; EMBL:Z99759; PIDN:CABL6903.1; GSPDB:GN00067; SPDB:SP1  
A;Experimental source: strain 972h-; cosmid c16E9  
C;Genetics:  
A;Gene: SPDB:SPBC16E9.11c  
A;Map position: 2  
A;Introns: 60/2  
C;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
F;236-273/Domain: WW repeat homology <WW1>  
F;306-343/Domain: WW repeat homology <WW2>  
F;364-401/Domain: WW repeat homology <WW3>

Query Match 59.6%; Score 130; DB 2; Length 786;  
Best Local Similarity 62.9%; Pred. No. 1.2e-08;  
Matches 22; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy 3 LPFGWEERQDILGRYYVNHESRRTQWKRPDQD 37  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 238 LPFGWEERADSLGRYYVDHNTTRTTWTRPASSIN 272

## RESULT 6

T49744  
probable ubiquitin-protein ligase [imported] - Neurospora crassa  
N;Alternate names: protein B24B19.160  
C;Species: Neurospora crassa  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C;Accession: T49744  
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25022  
A;Accession: T49744  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-815 <SCH>

A;Accession: T37545

45K WW domain-containing protein - human

A;Reference number: A56954; NCBIID:93504370; PMID:7782358  
A;Accession: E56954  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-472 <SD>  
A:Cross-references: UNIPROT:P46938; EMBL:X80508; NID:G517178; PID:G5171  
C;Comment: This protein binds to the SH3 domain of the Yes proto-oncogene product.  
C:Genetics:  
C:A Gene: YAP65  
C:Superfamily: yes-associated protein; WW repeat homology  
F;156-193/Domains: WW repeat homology <WW1>  
F;215-252/Domains: WW repeat homology <WW2>

Query Match            41.7%; Score 91; DB 2; Length 472;  
Best Local Similarity      45.2%; Pred. No. 0.0007;  
Matches 14; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy         2 PLPGWEERQDILGRYYVNHESRRQTOWKRP 32



Db 216 PLPDGWEQAMTQDGEVYVYINHKNKTTISWLDP 246

RESULT 15

I50730

Yes-associated protein, 65K - chicken

C:Species: Gallus gallus (chicken)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004

C:Accession: I50730; S39260

R:Sudo1, M.

Oncogene 9, 2145-2152, 1994

A:Title: Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to t

A:Reference number: I50730; MUID:94309887; PMID:8035999

A:Accession: I50730

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-448 <SUD>

A:Cross-references: UNIPROT:P46936; EMBL:X76483; NID:g434017; PIDN:CAA54021.1; PID:g4340

C:Genetics:

A:Gene: YAP65

C:Superfamily: yes-associated protein; WW repeat homology

F:169-206/Domain: WW repeat homology <WW1>

Query Match 40.4%; Score 88; DB 2; Length 448;

Best Local Similarity 45.2%; Pred. No. 0.0016;

Matches 14; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 2 PLPDGWEERQDILGRYYVNHESRRTQWKRP 32

||||||| : || : || : || : ||

Db 170 PLPDGWEAMKTPSGQRYFLNHIDQTTTWDQDP 200

Search completed: October 13, 2005, 15:11:14  
Job time : 15.8681 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 72.381 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-3  
Perfect score: 218  
Sequence: 1 SPLPGEWERDILGRYYVNHESRTQWKRPQDNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_crembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	218	100.0	1000	1 NED4_HUMAN	P46934 homo sapien
2	213	97.7	455	2 Q9N134	Q9N134 oryctolagus
3	203	93.1	887	1 NED4_MOUSE	P46935 mus musculu
4	199	91.3	887	1 NED4_RAT	Q62940 rattus norv
5	148	67.9	442	2 Q80U03	Q80U03 mus musculu
6	148	67.9	820	2 Q8NT88	Q8NT88 homo sapien
7	148	67.9	835	2 Q8BRT9	Q8BRT9 mus musculu
8	148	67.9	854	2 Q9H2W4	Q9H2W4 homo sapien
9	148	67.9	855	2 Q9BW58	Q9BW58 homo sapien
10	148	67.9	855	2 Q99PK2	Q99PK2 mus musculu
11	148	67.9	855	2 Q8CFI0	Q8CFI0 mus musculu
12	148	67.9	875	2 Q641N8	Q641N8 mus musculu
13	148	67.9	911	2 Q8NSA7	Q8NSA7 homo sapien
14	148	67.9	947	2 Q7Z5F1	Q7Z5F1 homo sapien
15	148	67.9	955	2 Q96PU5	Q96PU5 homo sapien
16	148	67.9	967	2 Q7Z5F2	Q7Z5F2 homo sapien
17	148	67.9	975	2 Q7Z5N3	Q7Z5N3 homo sapien
18	148	67.9	995	2 Q43165	Q43165 homo sapien
19	145	66.5	571	2 Q8GGJ2	Q8GGJ2 gallus gall
20	143	65.6	970	2 Q6DIR6	Q6DIR6 xenopus tro
21	142	65.1	971	2 Q42573	Q42573 xenopus lae
22	142	65.1	971	2 Q6GMD5	Q6GMD5 xenopus lae
23	139	63.8	917	2 Q7QE76	Q7QE76 anopheles g
24	138	63.3	793	2 Q7KUR2	Q7KUR2 drosophila
25	138	63.3	834	2 Q9SR64	Q9SR64 drosophila
26	138	63.3	838	2 Q95TQ0	Q95TQ0 drosophila
27	138	63.3	956	2 Q81QR6	Q81QR6 drosophila
28	138	63.3	1007	2 Q9VVI3	Q9VVI3 drosophila
29	135	61.9	1048	2 Q7PQR5	Q7PQR5 anopheles g
30	132	60.6	809	1 RSP5_YEAST	P39940 saccharomyc
31	132	60.6	817	2 Q75AI2	Q75AI2 ashbya goss

32 132 60.6 819 2 Q6CNC7 Q6cnc7 kluyveromyc  
33 132 60.6 822 2 Q6FN71 Q6fn71 candida gla  
34 131 60.1 781 2 Q6BT41 Q6bt41 debaryomyce  
35 130 59.6 786 1 PUB3\_SCHPO O14326 schizosacch  
36 130 59.6 787 2 Q6M906 Q6m906 neurospora  
37 130 59.6 806 2 Q7RV01 Q7rv01 neurospora  
38 128.5 58.9 870 1 WWP2\_HUMAN O0308 homo sapien  
39 128.5 58.9 870 1 WWP2\_MOUSE Q9DBH0 mus musculu  
40 128 58.7 806 2 Q6P289 Q6p289 mus musculu  
41 128 58.7 864 1 ITCH\_MOUSE Q8C863 mus musculu  
42 128 58.7 903 1 ITCH\_HUMAN Q9G102 homo sapien  
43 127 58.3 918 1 WWP1\_MOUSE O8BZZ3 mus musculu  
44 126 57.8 949 2 Q9YOH4 Q9yoh4 drosophila  
45 123 56.4 767 1 PUB1\_SCHPO Q92462 schizosacch

## ALIGNMENTS

RESULT 1  
NED4\_HUMAN  
ID NED4\_HUMAN STANDARD; PRT; 1000 AA.  
AC P46934;  
DT 01-~~Nov-1995~~ (Rel. 32, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).  
GN Name=NEDD4; Synonyms=KIAA0093;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP CONCEPTUAL TRANSLATION OF 1-73.  
RA Hinz U.;  
RL Unpublished observations (JUN-2003).  
RN [2]  
RP SEQUENCE OF 74-1000 FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=95308325; PubMed=7788527;  
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,  
RA Tabata S., Iihikawa K.-I., Kawarabayashi Y., Kohani H., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. III.  
RT analysis of cDNA clones from human cell line KG-1.";  
RL DNA Res. 2:37-43(1995).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
CC then directly transfers the ubiquitin to targeted substrates (By  
CC similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Interacts with UBE2D2. Binds SCNN1A, SCNN1B and SCNN1G  
CC (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
CC thiolester formation.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -!- SIMILARITY: Contains 4 WW domains.  
CC -!- CAUTION: The sequence of the N-terminus was deduced from the  
CC genomic sequence.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; AC039057; -; NOT ANNOTATED\_CDS.  
EMBL; D42055; BAA07655.1; -;

```

DR HSP; Q62940; 115H.
DR Genew: HGNC:7727; NEDD4.
DR H-InVDB; HIX0012269; -.
DR MIM; 602278; -.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2 CalB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS0020; WW_DOMAIN_2; 4.
KW Ligase; Repeat; Ub1 conjugation; Ub1 conjugation pathway.
FT DOMAIN 9 44 Pro-rich.
FT DOMAIN 77 81 Poly-Ser.
FT DOMAIN 106 207 C2 domain.
FT DOMAIN 285 288 Poly-Gln.
FT DOMAIN 291 324 WW 1.
FT DOMAIN 448 481 WW 2.
FT DOMAIN 521 554 WW 3.
FT DOMAIN 573 606 WW 4.
FT DOMAIN 665 1000 HECT.
FT BINDING 967 967 Ubiquitin (By similarity).
FT SEQUENCE 1000 AA; 114936 MW; 3728088E50C149CB CRC64;
Query Match 100.0%; Score 218; DB 1; Length 1000;
Best Local Similarity 100.0%; Pred. No. 7.8e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38
Db 291 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 328

RESULT 2
Q9N134 PRELIMINARY; PRT; 455 AA.
ID Q9N134
AC Q9N134;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-protein ligase Nedd4 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=21113801; PubMed=11181416;
RA Velazquez H., Silva T., Andujar E., Desir G.V., Ellison D.H.,
RA Greger R.
RT "The distal convoluted tubule of rabbit kidney does not express a
RT functional sodium channel.";
RL Am. J. Physiol. Renal Physiol. 280:F530-F539(2001).
DR EMBL; AF229024; AAF45194.1; -.
DR HSP; Q62940; 115H.
DR GO; GO:0016874; F:ligase activity; IEA.
DR InterPro; IPR008973; C2 CalB.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 4.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS0020; WW_DOMAIN_2; 4.
KW Ligase.
FT NON_TER 1

```

```

FT NON_TER 455
SQ SEQUENCE 455 AA; 51811 MW; D1C4549B5E22298 CRC64;

Query Match 97.7%; Score 213; DB 2; Length 455;
Best Local Similarity 97.4%; Pred. No. 1.5e-19;
Matches 37; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38
Db 98 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 135

RESULT 3
NED4_MOUSE
ID_NED4_MOUSE STANDARD; PRT; 887 AA.
AC P46935; O08758; O8GBB3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN Name=Nedd4; Synonyms=Kiaa0093, Nedd-4, Nedd4a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=92328780; PubMed=1378265;
RA Kumar S., Tomooka Y., Noda M.;
RT "Identification of a set of genes with developmentally down-regulated
RT expression in the mouse brain.";
RL Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
RN 2
RP REVISIONS.
RA Kumar S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

[3]
SEQUENCE FROM N.A. SUBCELLULAR LOCATION, AND INTERACTION WITH UBE2D2.
RC STRAIN=CS7BL/6 X CBA;
RX MEDLINE=97326076; PubMed=9182527; DOI=10.1074/jbc.272.24.15085;
RA Hatakeyama S., Jensen J.P., Weissman A.M.;
RT "Subcellular localization and ubiquitin-conjugating enzyme (E2)
RT interactions of mammalian HECT family ubiquitin protein ligases.";
RL J. Biol. Chem. 272:15085-15092(1997).
RN 4
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Nagataya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

```

FT	DOMAIN	249	282	WW 1.
FT	DOMAIN	405	438	WW 2.
FT	DOMAIN	460	493	WW 3.
FT	DOMAIN	552	887	HECT.
FT	BINDING	854	854	Ubiquitin (By similarity).
SQ	SEQUENCE	887 AA; 102705 MW; AE7DD3ED6396C50 CR64;		

Query Match 93.1%; Score 203; DB 1; Length 887;  
 Best Local Similarity 89.5%; Pred. No. 6.5e-18;  
 Matches 34; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	1	SPLPDQWEERQDILGRITYYVNHESRRRTQWKRPDPQDL 38
DB	249	SPLPDQWEERQDVLGRITYYVNHESRRRTQWKRPSPDDDL 286

RESULT 4  
 NED4 RAT STANDARD; PRT; 887 AA.  
 AC Q62940;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).  
 GN Names=Nedd4;  
 OS Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.  
 RC TISSUE=Lung;  
 RX MEDLINE=96221297; PubMed=8665844;  
 RA Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,  
 RA Robin D.;  
 RT "WW domains of Nedd4 bind to the proline-rich PY motifs in the  
 RT epithelial Na+ channel deleted in Liddle's syndrome.";  
 RL EMBO J. 15:2371-2380(1996).  
 RN [2]  
 RP STRUCTURE BY NMR OF 452-499 IN COMPLEX WITH SCNN1B, AND INTERACTION  
 RP WITH SCNN1A; SCNN1B AND SCNN1G.  
 RX MEDLINE=21223577; PubMed=11323714; DOI=10.1038/87562;  
 RA Kanelis V., Rotin D., Forman-Kay J.D.;  
 RT "Solution structure of a Nedd4 WW domain-ENaC peptide complex.";  
 RL Nat. Struct. Biol. 8:407-412(2001).  
 CC -1- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates (By  
 CC similarity).  
 CC -1- SUBUNIT: Ubiquitin conjugation; third step.  
 CC -1- SUBUNIT: Interacts with UBE2D2 (By similarity). Binds SCNN1A,  
 CC SCNN1B and SCNN1G.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Expression is highest in lung,  
 CC kidney and brain.  
 CC -1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -1- SIMILARITY: Contains 1 C2 domain.  
 CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -1- SIMILARITY: Contains 3 WW domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U50842; AAB48949.1; -.  
 CC DR PIR; S70642; S70642.  
 CC DR PDB; 1ISH; NMR; W=450-499.  
 CC DR



RA	Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AK042621; BAC31307.1; -
DR	HSSP; Q62940; 115H.
DR	MGI; MGI:1933754; Nedd41.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InterPro; IPR000569; HECT.
DR	InterPro; IPR002349; WW.
DR	InterPro; IPR001202; WW_Rsp5_WWP.
DR	Pfam; PF00632; HECT; 1.
DR	Pfam; PF00397; WW; 4.
DR	PRINTS; PR00403; WWDOMAIN.
DR	SMART; SM00119; HECTc; 1.
DR	SMART; SM00456; WW; 4.
DR	PROSITE; PS05237; HECT; 1.
DR	PROSITE; PS01159; WW_DOMAIN1; 4.
DR	PROSITE; PS05020; WW_DOMAIN2; 4.
DR	SEQUENCE 835 AA; 96410 MW; 81c7DD62722DDA9B CRC64;
Query Match	67.9%; Score 148; DB 2; Length 835;
Best Local Similarity	71.4%; Pred. No. 1.1e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gap	
QY	2 PLPPGWEERQDILGRYYVNNHESRRRTQWKRTPOD 36
	: : : : : : : : : : : : : : : : : :
DB	73 PLPPGWEKVDNLGRYYVNNHNRSTQWRPSLMD 107
RESULT 8	
Q9H2W4	PRELIMINARY; PRT; 854 AA.
ID	Q9H2W4
AC	Q9H2W4;
DT	01-11-2001 (TrEMBLrel. 16, Created)
DT	01-11-2001 (TrEMBLrel. 16, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	NEDD4La (Ubiquitin ligase NEDD4Lb) (Ubiquitin ligase NEDD4F).
GN	Name=NEDD4La; Synonyms=NEDD4L;
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
OX	[1]
RP	SEQUENCE FROM N.A.
RP	MEDLINE=21928840; PubMed=11840194; DOI=10.1038/sj.ejhg.5200747;
RA	Chen H., Ross C.A., Wang N., Huo Y., MacKinnon D.F., Pozash J.B.,
RA	Simpson S.G., McMahon F.J., DePaulo J.R. Jr., McInnis M.G.;
RT	"NEDD4L on human chromosome 19q21 has multiple forms of transcripts
RT	and is a homologue of the mouse Nedd4-2 gene.";
RT	Eur. J. Hum. Genet. 9:922-930(2001).
RL	[2]
RP	SEQUENCE FROM N.A.
RP	MEDLINE=22980412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009;
RX	Qi H., Grenier J., Fournier A., Labrie C.;
RT	"Androgens differentially regulate the expression of NEDD4L
RT	transcripts in LNCaP human prostate cancer cells.";
RL	Mol. Cell. Endocrinol. 210:51-62(2003).
RL	EMBL; AF210730; AAG43524.1; -
DR	EMBL; AF385931; AAM46208.1; -
DR	EMBL; AF112983; AAM76728.1; -
DR	HSSP; Q62940; 115H.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0015874; F:ligase activity; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InterPro; IPR000569; HECT.
DR	InterPro; IPR002349; WW.
DR	InterPro; IPR001202; WW_Rsp5_WWP.
DR	Pfam; PF00632; HECT; 1.
DR	Pfam; PF00397; WW; 4.
DR	PRINTS; PR00403; WWDOMAIN.
DR	SMART; SM00119; HECTc; 1.





Q8CF10	PRELIMINARY;	PRT;	855 AA.
AC	Q8CF10		
AD	Q8CF10		
DT	01-WAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Neural cell expressed, developmentally down-regulated gene 4-like.		
DE	Name=Nedd41;		
GN	Mus musculus (Mouse).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=FVB/N; TISSUE=Kidney, and Mammary tumor. C3;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RC	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Goughin J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Trinchwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;		
RC	Strausberg R.;		
RA	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=FVB/N; TISSUE=Kidney;		
RC	Strausberg R.;		
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC039746; AAH39746.1; -		
DR	EMBL; BC071210; AAH71210.1; -		
DR	HSSP; Q62940; 115H.		
DR	MGI; 1933754; Nedd41.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.		
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.		
DR	InterPro; IPR000569; HECT.		
DR	InterPro; IPR002349; WW.		
DR	InterPro; IPR001202; WW_Rsp5_WWP.		
DR	Pfam; PF00632; HECT; 1.		
DR	Pfam; PF00397; WW; 4.		
DR	PRINTS; PR00403; WDDOMAIN.		
DR	SMART; SM00119; HECTC; 1.		
DR	SMART; SM00456; WW; 4.		
DR	PROSITE; PS02037; HECT; 1.		
DR	PROSITE; PS01159; WW_DOMAIN_1; 4.		
DR	PROSITE; PS50020; WW_DOMAIN2; 4.		
DR	SEQUENCE 855 AA; 98465 MW; 96C4528442855895 CRC64;		
SQL			
Query Match	67.9%;	Score 148;	DB 2; Length 855;
Best Local Similarity	71.4%;	Pred No. 1.1e-10;	
Matches	25; Conservative	3; Mismatches	7; Indels 0; Gaps
QY	2 PLPQWEERQDILGRYYVNHESRRTWKRPD 36		

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 C2 domain.
DR ENBL; BC032597; AAI32597.1; -.
DR HSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
KW Ligase.
SQ SEQUENCE 947 AA; 109429 MW; 95CB0FDEDCB96639 CRC64;

Query Match 67.9%; Score 148; DB 2; Length 947;
Best Local Similarity 71.4%; Pred. No. 1.2e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PLPPGWEERODILGRYYVNHESRRTQWKRTPOD 36
Db 194 PLPPGWEKVDNLGRYYVNHNNRTTQWHRPSLMD 228

RESULT 14
ID Q7Z5F1 PRELIMINARY; PRT; 947 AA.
AC Q7Z5F1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin ligase NEDD4h.
GN Name=NEDD4L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20501262; PubMed=11046148;
RX DOI=10.1128/MCB.20.22.8526-8535.2000;
RA Winberg G., Mateskova L., Chen F., Plant P., Rotin D., Gish G.,
RA Ingham R., Ernberg I., Pawson T.;
RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RL protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";
RL Mol. Cell. Biol. 20:8526-8535(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269431; PubMed=11244092; DOI=10.1074/jbc.C000906200;
RX Harvey F.K., Dinudom A., Cook I.D., Kumar S.;
RT "The Nedd4-like protein KIAA0439 is a potential regulator of the
RT epithelial sodium channel.";
RL J. Biol. Chem. 276:8597-8601(2001).
RN [3]
RP SEQUENCE FROM N.A.
```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009;
RA Qi H., Grenier J., Fournier A., Labrie C.;
RT "Androgens differentially regulate the expression of NEDD4L
RT transcripts in LNCap human prostate cancer cells.";
RL Mol. Cell. Endocrinol. 210:51-62(2003).
CC -1- SIMILARITY: Contains 1 C2 domain.
DR ENBL; AY112985; AAM76730.1; -.
DR HSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
KW Ligase.
SQ SEQUENCE 947 AA; 109429 MW; 95CB0FDEDCB96639 CRC64;

Query Match 67.9%; Score 148; DB 2; Length 947;
Best Local Similarity 71.4%; Pred. No. 1.3e-10;
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PLPPGWEERODILGRYYVNHESRRTQWKRTPOD 36
Db 186 PLPPGWEKVDNLGRYYVNHNNRTTQWHRPSLMD 220

RESULT 15
ID Q96PU5 PRELIMINARY; PRT; 955 AA.
AC Q96PU5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NEDD4-like ubiquitin ligase 3.
GN Name=NEDL3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20501262; PubMed=11046148;
RX DOI=10.1128/MCB.20.22.8526-8535.2000;
RA Winberg G., Mateskova L., Chen F., Plant P., Rotin D., Gish G.,
RA Ingham R., Ernberg I., Pawson T.;
RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RL protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";
RL Mol. Cell. Biol. 20:8526-8535(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269431; PubMed=11244092; DOI=10.1074/jbc.C000906200;
RX Harvey F.K., Dinudom A., Cook I.D., Kumar S.;
RT "The Nedd4-like protein KIAA0439 is a potential regulator of the
RT epithelial sodium channel.";
RL J. Biol. Chem. 276:8597-8601(2001).
RN [3]
RP SEQUENCE FROM N.A.
```

RA Okamoto Y., Miyazaki K., Sakamoto M., Kato C., Nakagawara A.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
DR EMBL; AB071179; BAB69424.1; -.  
DR HSSP; Q62940; 115H.  
DR GO; GO:0005622; C:intracellular; IC.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR GO; GO:0017080; F:sodium channel regulator activity; IDA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.  
DR GO; GO:0007588; P:excretion; NAS.  
DR GO; GO:0045807; P:positive regulation of endocytosis; NAS.  
DR GO; GO:0016567; P:protein ubiquitination; NAS.  
DR GO; GO:0042176; P:regulation of protein catabolism; NAS.  
DR GO; GO:0010038; P:response to metal ion; IDA.  
DR GO; GO:0006883; P:sodium ion homeostasis; NAS.  
DR GO; GO:0006814; P:sodium ion transport; NAS.  
DR GO; GO:0030104; P:water homeostasis; NAS.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CaLB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00360; C2DOMAIN.  
DR PRINTS; PR00403; WDDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 4.  
KW Ligase.  
SQ SEQUENCE 955 AA; 110021 MW; A8BB278A37F6A6B5 CRC64;

Query Match 67.9%; Score 148; DB 2; Length 955;  
Best Local Similarity 71.4%; Pred. No. 1.3e-10;  
Matches 25; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PLPGWEERQDILGRYYVNHESRRTQWKRTPOD 36  
|||||:|||||:|||||:|||||:  
Db 194 PLPGWEKVDNLGRYYVNHNRRTQWHRPSLMD 228

Search completed: October 13, 2005, 15:09:16  
Job time : 73.381 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 85.6044 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-4  
Perfect score: 217  
Sequence: 1 SGLPPGWEERDILGRYYVNHESRTQWKRPDQNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	100.0	38	AAB83019	Aab83019 Human Ned
2	209	96.3	38	AAB83018	Aab83018 Human Ned
3	209	96.3	733	AAE32726	Aae32726 Human NED
4	209	96.3	900	ADQ07991	Adq07991 Human NED
5	209	96.3	927	AAy25170	Aay25170 Human KIA
6	209	96.3	927	AAE32725	Aae32725 Human NED
7	200	92.2	38	ADB49212	Adb49212 Mouse WW
8	194	89.4	38	AAR97680	Aar97680 Mouse Ned
9	194	89.4	38	AAAB21978	Aab21978 Nedda/mou
10	194	89.4	38	ADB49211	Adb49211 Human WW
11	146	67.3	38	AAR97681	Aar97681 Mouse Ned
12	146	67.3	38	ADB49214	Adb49214 Mouse WW
13	141	65.0	38	ADB49215	Adb49215 Human WW
14	141	65.0	818	AAW13386	Aaw13386 Human pro
15	141	65.0	830	ADQ97134	Adq97134 Human can
16	141	65.0	834	AAU77715	Aau77715 Diseased
17	141	65.0	854	ABP58332	Abp58332 Human cel
18	141	65.0	854	ADB75453	Adb75453 Prostate
19	141	65.0	911	ADB75455	Adb75455 Prostate
20	141	65.0	923	ABG01080	Abg01080 Novel hum
21	141	65.0	940	ADR66480	Adr66480 Human pro
22	141	65.0	940	ADR66138	Adr66138 Human pro
23	141	65.0	941	ADQ97130	Adq97130 Human can
24	141	65.0	948	ADO55161	Ado55161 Protein #
25	141	65.0	954	ADQ97125	Adq97125 Mouse can

26	141	65.0	955	8	ADQ07990	Adq07990 Human. NED
27	141	65.0	958	8	ADQ97128	Adq97128 Human can
28	141	65.0	975	2	AAW93167	Aaw93167 Human ZGC
29	141	65.0	995	6	AAE32719	Aae32719 KIAA0439
30	141	65.0	995	7	ADB75451	Adb75451 Prostate
31	141	65.0	995	7	ADP60011	Adp60011 Human pro
32	141	65.0	995	7	ADD45762	Add45762 Human pro
33	141	65.0	995	8	ADR04136	Adr04136 Human Ned
34	141	65.0	1000	8	ADQ97132	Adq97132 Human can
35	141	65.0	1071	4	ABG11772	Abg11772 Novel hum
36	138	63.6	1082	4	ABB63355	Abb63355 Drosophil
37	135	62.2	37	2	AAW37638	Aaw37638 Repide c
38	135	62.2	38	7	ADB49230	Adb49230 Novel hum
39	133	61.3	38	5	AAU87972	Aau87972 Human WW
40	133	61.3	38	7	ADB49326	Adb49326 Novel hum
41	133	61.3	724	2	AAW36797	Aaw36797 Novel hum
42	133	61.3	725	7	ADB49320	Adb49320 Novel hum
43	132	60.8	38	2	AAR97683	Aar97683 Yeast Rep
44	132	60.8	38	3	AAAB21979	Aab21979 RSPS/9C p
45	132	60.8	38	4	AAAB83022	Aab83022 Yeast Rep

ALIGNMENTS

RESULT 1  
AAB83019  
ID AAB83019 standard; peptide; 38 AA.

AC AAB83019;  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Human Nedda HECT E3 ubiquitin ligase WW domain #2.

KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytosolic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; WW domain; Nedda.  
XX Homo sapiens.

XX WO200116604-A1.  
XX  
XX  
PD 08-MAR-2001.

PF 29-AUG-2000; 2000WO-US023729.

PR 30-AUG-1999; 99US-00385918.

PA (SIGN-) SIGNAL PHARM INC.

PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;

WPI; 2001-327913/34.

PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
PT evaluating the ability of agents to modulate Smad protein degradation.

XX Disclosure; Page 13; 75pp; English.

XX The present sequence is the WW domain of a HECT (homologous to E6  
CC carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad  
CC PY motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.  
CC The sequence is provided in a specification relating to a method for  
CC screening for agents that modulate transforming growth factor (TGF)-beta  
CC and/or bone morphogenic protein (BMP)-mediated signalling. The method  
CC involves evaluating the effect of an agent on binding of HECT E3  
CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein

CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
CC bone formation in a patient or treating a condition associated with  
CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
CC inhibit BMP-mediated signalling are useful for treating inflammation,  
CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
CC signalling are useful for stimulating bone anabolism as well as treating  
CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
CC that inhibit TGF-mediated signalling are useful for treating cancer,  
CC inflammation, neurodegeneration and fibrosis  
XX Sequence 38 AA;  
SQ Query Match 100.0%; Score 217; DB 4; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.2e-21;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
Db 1 SGLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38

RESULT 2  
AAB83018  
ID AAB83018 standard; peptide; 38 AA.  
XX  
AC AAB83018;  
XX  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Human Nedd4 HECT E3 ubiquitin ligase WW domain #1.  
XX  
KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; WW domain; Nedd4.  
XX  
OS Homo sapiens.  
XX  
XX WO200116604-A1.  
XX  
PD 08-MAR-2001.  
XX  
XX 29-AUG-2000; 2000WO-US023729.  
XX  
XX 30-AUG-1999; 99US-00385918.  
XX  
XX (SIGN-) SIGNAL PHARM INC.  
XX  
XX Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
XX  
XX WPI; 2001-327913/34.  
XX  
XX Screening for modulators of TGF-beta and/or bone morphogenic protein  
XX (BMP) mediated signalling useful for treating cancer and osteoporosis by  
XX evaluating the ability of agents to modulate Smad protein degradation.  
XX  
XX Disclosure; Page 13; 75pp; English.

CC The present sequence is the WW domain of a HECT (homologous to E6  
CC carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad  
CC PY motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.  
CC The sequence is provided in a specification relating to a method for  
CC screening for agents that modulate transforming growth factor (TGF)-beta  
CC and/or bone morphogenic protein (BMP)-mediated signalling. The method  
CC involves evaluating the effect of an agent on binding of HECT E3  
CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein  
CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
CC bone formation in a patient or treating a condition associated with  
CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
CC inhibit BMP-mediated signalling are useful for treating inflammation,

CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
CC signalling are useful for stimulating bone anabolism as well as treating  
CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
CC that inhibit TGF-mediated signalling are useful for treating cancer,  
CC inflammation, neurodegeneration and fibrosis  
XX Sequence 38 AA;  
SQ Query Match 96.3%; Score 209; DB 4; Length 38;  
Best Local Similarity 97.4%; Pred. No. 1.4e-20;  
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38  
Db 1 SGLPPGWEERQDILGRYYVYNHESRRRTQWKRPDPQDNL 38

RESULT 3  
AAE32726  
ID AAE32726 standard; protein; 733 AA.  
XX  
AC AAE32726;  
XX  
XX 24-MAR-2003 (first entry)  
XX  
XX Human NEDD4 short form protein.  
XX  
XX Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;  
KW poliomyelitis; HIV; measles; protein therapy; human; NEDD4.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 260 /note= "Encoded by CAA"  
FT Misc-difference 279 /note= "Encoded by AGC"  
XX  
XX WO200290549-A2.  
XX  
XX 14-NOV-2002.  
XX  
XX 12-MAR-2002; 2002WO-1B002106.  
XX  
XX 12-MAR-2001; 2001US-0275224P.  
XX 31-JUL-2001; 2001US-0308958P.  
XX 07-DEC-2001; 2001US-0340170P.  
XX  
XX (PROT-) PROTEOLOGICS LTD.  
XX  
XX Greener T, Moskowitz H, Reiss Y, Alroy I;  
XX  
XX WPI; 2003-111976/10.  
XX N-PSDB; AAD50460.  
XX  
XX New protein complex comprising HECT-RCC1, viral maturation scaffolding  
XX protein (VMSp) and/or HIV gag protein, useful for treating viral  
XX infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,  
XX or Ebola.  
XX  
XX Disclosure; Fig 16; 150pp; English.

CC The invention relates to a method for modulation of viral maturation. The  
XX invention also provides an isolated protein complex comprising a HECT-  
CC RCC1 selected from HECT-MW, HECT-RCC1, Gag protein, Gag late domain, P13,  
CC actin, myosin, Hsp60, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIIM,  
CC GTPase, E2 enzyme, tsgl01, cullin, HEC1, HEC2, HEC3, Nedd4 -like  
CC protein or clathrin. The complexes, proteins, antibodies and methods are  
CC useful for treating viral infections, such as lymphosarcoma, human  
CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola  
CC and for inhibiting budding in a subject. They are also useful in  
CC diagnostic assays for determining whether a cell is infected with a virus  
CC and for characterising the nature, progression and/or infectivity of the







PT discovery, modification and refinement, for discovering polypeptides  
 PT involved in pharmacological activities, or as an immunogen to generate  
 PT antibodies.

XX Example; Fig 5; 133pp; English.

XX The invention describes a purified polypeptide (I) comprising a WW domain  
 CC which has a sequence (S1) selected from 11 sequences fully defined in the  
 CC specification, a sequence (S2) selected from 48 sequences fully defined in the  
 CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725  
 CC amino acids fully defined in the specification. (I) is useful for  
 CC screening a potential drug candidate, by allowing (I) to come into  
 CC contact with at least one recognition unit having a selective affinity  
 CC for the WW domain in (I), in the presence of an amount of a potential  
 CC drug candidate, such that (I) and the recognition unit are capable of  
 CC interacting when brought into contact with one another in the absence of  
 CC the drug candidate, and determining the effect, if any, of the presence  
 CC of the amount of the drug candidate on the interaction of (I) with the  
 CC recognition unit. (I) is useful for drug discovery, modification and  
 CC refinement, for discovering polypeptides involved in pharmacological  
 CC activities, or as an immunogen to generate antibodies. This is the amino  
 CC acid sequence of WW binding protein Ned4 WW domain.

XX Sequence 38 AA;

Query Match 92.2%; Score 200; DB 7; Length 38;  
 Best Local Similarity 94.7%; Pred. No. 2.3e-19;  
 Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38  
 Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38

RESULT 8

ID AAR97680 standard; protein; 38 AA.

XX AAR97680;

XX 30-AUG-1996 (first entry)

XX Mouse Nedd4 WW domain-1.

XX WW domain; signal transduction; diagnosis; gene therapy;  
 KW Yes proto-oncogene associated protein; YAP; Nedd4.

XX Mus sp.

XX WO9617061-A1.

XX 06-JUN-1996.

XX 30-NOV-1995; 95WO-US015512.

XX 01-DEC-1994; 94US-00348518.

XX 07-JUN-1995; 95US-00476509.

XX (UVRQ) UNIV ROCKEFELLER.

XX (DELB-) DELBRUECK CENT MOLECULAR MEDICINE MAX.

XX Sudol M, Chen H, Bork P;

XX WPI; 1996-286829/29.

XX DNA encoding Yes proto-oncogene associated protein - used to modulate

XX intracellular signal transduction e.g. for treatment of muscular

XX dystrophy.

XX Claim 10; Fig 12; 126pp; English.

XX WW domains (AAR97673-92) were identified in a number of proteins involved

CC in signalling or regulatory functions. The WW domain was initially

CC identified in the Yes proto-oncogene associated proteins (YAPs) of  
 CC chicken, human and mouse (see also AAR97669-70 and AAR97672). A consensus  
 CC sequence is given in AAR97671. Mouse Nedd4 plays a role in embryonic  
 CC development and CNS differentiation. The Nedd4 WW domains (AAR97680-82)  
 CC can be expressed in bacterial, yeast, insect or mammalian cells, and used  
 CC to identify WW domain ligands. They can be introduced into cells, either  
 CC directly or by gene therapy, to increase the level of signal transduction

XX Sequence 38 AA;

Query Match 89.4%; Score 194; DB 2; Length 38;

Best Local Similarity 86.8%; Pred. No. 1.5e-18;

Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQDNL 38

Db 1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPSPDDDL 38

RESULT 9

ID AAB21978 standard; peptide; 38 AA.

XX AAB21978;

XX 02-JAN-2001 (first entry)

XX Nedd4/mouse peptide containing a WW-domain #2.

XX WW-domain; protein-protein interaction; cell growth regulation;  
 KW protein degradation regulation; Alzheimer's; Dementia pugilistica;  
 KW Down's syndrome; Parkinson's disease; Pick's; neurodegenerative;  
 KW microtubule assembly; tau; hyperplasia; neoplasia; psoriasis;  
 KW retinosis; atherosclerosis; leukaemia; lymphoma; papiloma;  
 KW pulmonary fibrosis; rheumatoid arthritis; multiple sclerosis;  
 KW muscular dystrophy; mouse.

XX Mus sp.

XX WO200048621-A2.

XX 24-AUG-2000.

XX 18-FEB-2000; 2000WO-US004278.

XX 18-FEB-1999; 99US-00252404.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Lu KP, Zhou XZ;

XX WPI; 2000-594014/56.

XX Mediating protein-protein interactions, useful for regulating cell growth  
 PT and for treating neurodegenerative disorders, comprises modulating  
 PT binding of WW domain containing polypeptide with phosphorylated ligand.

XX Disclosure; Fig 6; 82pp; English.

XX The present invention relates to a method for mediating protein-protein  
 CC interaction, which comprises modulating the binding of a WW-domain  
 CC containing peptide with a phosphorylated ligand e.g. tau. WW-domains are  
 CC highly conserved regions of approximately 40 amino acid residues with two  
 CC invariant tryptophans (W) in a triple stranded beta-sheet. The present  
 CC sequence is one such WW-domain. When a WW-domain containing peptide is  
 CC phosphorylated at serine or threonine residues, dephosphorylation of  
 CC ligands bound to the peptide is inhibited. The present peptide may be  
 CC useful for mediating protein-protein interaction, regulating cell growth,  
 CC regulating protein degradation, restoring the function of tau to bind  
 CC microtubules and promote or restore microtubule assembly in  
 CC neurodegenerative diseases e.g. Alzheimer's, Dementia pugilistica, Down's  
 CC syndrome, Parkinson's disease, Pick's disease, multiple sclerosis,  
 CC muscular dystrophy, Corticobasal degeneration, Frontotemporal dementias,

CC Myotonic dystrophy, Niemann-Pick disease, prion disease with tangles,  
CC progressive supranuclear palsy and subacute sclerosing panencephalitis.  
CC In addition, inhibitors or stimulators of interactions between WW-domains  
CC and ligands of the present invention can be used to treat hyperplastic  
CC and neoplastic disorders e.g. all forms of malignancies, psoriasis,  
CC retinosis, atherosclerosis resulting from plaque formation, leukaemias,  
CC benign tumour growth, lymphomas, papillomas, pulmonary fibrosis and  
CC rheumatoid arthritis  
XX  
XX

QQ Sequence 38 AA;

Query March 89.4%; Score 194; DB 3; Length 38;  
Best Local Similarity 86.8%; Pred. No. 1.5e-18;  
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLPPGWEERODILGRITYVNHESRRRTQWKRPDPDNL 38  
DB 1 SPLPPGWEERODVLGRITYVNHESRRRTQWKRPSPDDLL 38

RESULT 10

ADB49211

ID ADB49211 standard; peptide; 38 AA.

XX AC ADB49211;

XX XX

DT 04-DEC-2003 (first entry)

XX XX

DE Human WW binding protein Ned4 WW domain #1.

XX KW WW domain; drug candidate screening; drug discovery; drug modification;  
KW drug refinement; immunogen; WW binding protein; human; ned4; WW domain.  
XX OS Homo sapiens.  
XX PN US2003077577-A1.  
XX PD 24-APR-2003.  
XX PF 28-JUN-2002; 2002US-00185050.  
XX PR 03-APR-1996; 96US-00630916.  
PR 03-APR-1997; 97US-00826516.  
XX (PIRO/) PIROZZI G.  
PA (KAYB/) KAY B K.  
PA (FOWL/) FOWLKES D M.  
XX  
XX PI Pirozzi G, Kay BK, Fowlkes DM;  
XX WPI; 2003-635075/60.  
XX  
XX Novel purified polypeptide comprising WW domain, useful for drug  
FT discovery, modification and refinement, for discovering polypeptides  
FT involved in pharmacological activities, or as an immunogen to generate  
FT antibodies.  
XX  
XX Example; Fig 5; 133pp; English.

XX

CC The invention describes a purified polypeptide (I) comprising a WW domain  
CC which has a sequence (S1) selected from 11 sequences fully defined in the  
CC specification, a sequence (S2) selected from 48 sequences fully defined  
CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725  
CC amino acids fully defined in the specification. (I) is useful for  
CC screening a potential drug candidate, by allowing (I) to come into  
CC contact with at least one recognition unit having a selective affinity  
CC for the WW domain in (I), in the presence of an amount of a potential  
CC drug candidate, such that (I) and the recognition unit are capable of  
CC interacting when brought into contact with one another in the absence of  
CC the drug candidate, and determining the effect, if any, of the presence  
CC of the amount of the drug candidate on the interaction of (I) with the  
CC recognition unit. (I) is useful for drug discovery, modification and  
CC refinement, for discovering polypeptides involved in pharmacological

AD849214		Human WW binding protein Ned4 WW domain #2.
ID	ADB49214 standard; peptide; 38 AA.	
XX		
AC	ADB49214;	WW domain; drug candidate screening; drug discovery; drug modification;
KW		drug refinement; immunogen; WW binding protein; human; ned4; WW domain.
DT	04-DEC-2003 (first entry)	
XX		Homo sapiens.
DE	Mouse WW binding protein Ned4 WW domain #2.	
XX		
KW	WW domain; drug candidate screening; drug discovery; drug modification;	
KW	drug refinement; immunogen; WW binding protein; mouse; ned4; WW domain.	
XX		
OS	Mus sp.	
XX		
PN	US2003077577-A1.	
XX		
PD	24-APR-2003.	
XX		
PF	28-JUN-2002; 2002US-00185050.	
XX		
PR	03-APR-1996; 96US-00630916.	
PR	03-APR-1997; 97US-00826516.	
XX		(PIRO/) PIROZZI G.
PA	(KAYB/) KAY B K.	
PA	(FOWL/) FOWLKES D M.	
XX		Pirozzi G, Kay BK, Fowlkes DM;
PI		WPI; 2003-635075/60.
DR		
XX		Novel purified polypeptide comprising WW domain, useful for drug
PT		discovery, modification and refinement, for discovering polypeptides
PT		involved in pharmacological activities, or as an immunogen to generate
PT		antibodies.
XX		
PS	Example; Fig 5; 133pp; English.	
XX		The invention describes a purified polypeptide (I) comprising a WW domain
CC		which has a sequence (S1) selected from 11 sequences fully defined in the
CC		specification, a sequence (S2) selected from 48 sequences fully defined
CC		in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC		amino acids fully defined in the specification. (I) is useful for
CC		screening a potential drug candidate, by allowing (I) to come into
CC		contact with at least one recognition unit having a selective affinity
CC		for the WW domain in (I), in the presence of an amount of a potential
CC		interacting when brought into contact with one another in the absence of
CC		the drug candidate, and determining the effect, if any, of the presence
CC		of the amount of the drug candidate on the interaction of (I) with the
CC		recognition unit. (I) is useful for drug discovery, modification and
CC		refinement, for discovering polypeptides involved in pharmacological
CC		activities, or as an immunogen to generate antibodies. This is the amino
CC		acid sequence of WW binding protein Ned4 WW domain.
XX		
SQ	Sequence 38 AA;	
	Query Match 67.3%; Score 146; DB 7; Length 38;	
	Best Local Similarity 64.9%; Pred. No. 4.3e-12; Indels 0; Gaps 0;	
	Matches 24; Conservative 6; Mismatches 7; Indels 0; Gaps 0;	
Qy	1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPDN 37	
Db	1 SGLPPGWEERQDILGRYYVNHESRRRTTWSKTMDDD 37	
RESULT 13		
ADB49215		
ID	ADB49215 standard; peptide; 38 AA.	
XX		
AC	ADB49215;	
XX		
TT	04-DEC-2003 (first entry)	

XX		Human WW binding protein Ned4 WW domain #2.
DE		
XX		
KW	WW domain; drug candidate screening; drug discovery; drug modification;	
KW	drug refinement; immunogen; WW binding protein; human; ned4; WW domain.	
OS	Homo sapiens.	
XX		
PN	US2003077577-A1.	
XX		
PD	24-APR-2003.	
XX		
PF	28-JUN-2002; 2002US-00185050.	
XX		
PR	03-APR-1996; 96US-00630916.	
PR	03-APR-1997; 97US-00826516.	
XX		(PIRO/) PIROZZI G.
PA	(KAYB/) KAY B K.	
PA	(FOWL/) FOWLKES D M.	
XX		Pirozzi G, Kay BK, Fowlkes DM;
PI		WPI; 2003-635075/60.
DR		
XX		Novel purified polypeptide comprising WW domain, useful for drug
PT		discovery, modification and refinement, for discovering polypeptides
PT		involved in pharmacological activities, or as an immunogen to generate
PT		antibodies.
XX		
PS	Example; Fig 5; 133pp; English.	
XX		The invention describes a purified polypeptide (I) comprising a WW domain
CC		which has a sequence (S1) selected from 11 sequences fully defined in the
CC		specification, a sequence (S2) selected from 48 sequences fully defined
CC		in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC		amino acids fully defined in the specification. (I) is useful for
CC		screening a potential drug candidate, by allowing (I) to come into
CC		contact with at least one recognition unit having a selective affinity
CC		for the WW domain in (I), in the presence of an amount of a potential
CC		drug candidate, such that (I) and the recognition unit are capable of
CC		interacting when brought into contact with one another in the absence of
CC		the drug candidate, and determining the effect, if any, of the presence
CC		of the amount of the drug candidate on the interaction of (I) with the
CC		recognition unit. (I) is useful for drug discovery, modification and
CC		refinement, for discovering polypeptides involved in pharmacological
CC		activities, or as an immunogen to generate antibodies. This is the amino
CC		acid sequence of WW binding protein Ned4 WW domain.
XX		
SQ	Sequence 38 AA;	
	Query Match 65.0%; Score 141; DB 7; Length 38;	
	Best Local Similarity 68.6%; Pred. No. 2e-11; Indels 0; Gaps 0;	
	Matches 24; Conservative 4; Mismatches 7; Indels 0; Gaps 0;	
Qy	1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQ 35	
Db	1 SGLPPGWEERQDILGRYYVNHESRRRTTWTQVTQ 35	
RESULT 14		
AAW13386		
ID	AAW13386 standard; protein; 818 AA.	
XX		
AC	AAW13386;	
XX		
DT	10-JUL-1997 (first entry)	
XX		
DE	Human protein ubiquitin ligase pub3.	
XX		
KW	Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase; p53;	
KW	cell cycle; transgenic animal.	
XX		



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 18.0952 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-4  
Perfect score: 217  
Sequence: 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209	95.3	927	3	US-08-895-601-6
2	200	92.2	38	3	US-08-630-916A-18
3	194	89.4	38	2	US-09-066-074-11
4	194	89.4	38	2	US-08-555-912A-11
5	194	89.4	38	3	US-08-630-916A-17
6	194	89.4	38	3	US-08-348-518C-13
7	194	89.4	38	3	US-08-476-509B-13
8	194	89.4	38	4	US-09-252-404A-36
9	194	89.4	38	4	US-09-275-900-11
10	146	67.3	38	3	US-08-630-916A-20
11	146	67.3	38	3	US-08-348-518C-14
12	146	67.3	38	3	US-08-476-509B-14
13	141	65.0	38	3	US-08-630-916A-21
14	141	65.0	834	3	US-08-539-205A-6
15	141	65.0	834	4	US-09-392-163A-6
16	135	62.2	38	3	US-08-630-916A-36
17	132	60.8	38	2	US-09-066-074-12
18	132	60.8	38	2	US-08-555-912A-12
19	132	60.8	38	3	US-08-630-916A-19
20	132	60.8	38	3	US-08-348-518C-16
21	132	60.8	38	3	US-08-476-509B-16
22	132	60.8	38	4	US-09-252-404A-37
23	132	60.8	38	4	US-09-275-900-12
24	126	58.1	89	4	US-09-248-796A-18186
25	123	56.7	766	3	US-08-539-205A-4
26	123	56.7	766	4	US-09-392-163A-4
27	122	56.2	683	3	US-08-630-916A-46

28	121.5	56.0	906	3	US-08-630-916A-48	Sequence 48, Appl
29	121	55.8	30	4	US-09-252-404A-3	Sequence 3, Appl
30	121	55.8	38	3	US-08-630-916A-35	Sequence 35, Appl
31	121	55.8	474	4	US-09-774-639-371	Sequence 371, App
32	121	55.8	752	4	US-09-919-039-235	Sequence 235, App
33	121	55.8	852	2	US-09-070-060-3	Sequence 3, Appl
34	121	55.8	852	3	US-09-357-746-3	Sequence 3, Appl
35	121	55.8	854	2	US-09-070-060-4	Sequence 4, Appl
36	121	55.8	854	3	US-09-357-746-4	Sequence 4, Appl
37	112	51.6	38	3	US-08-630-916A-25	Sequence 25, Appl
38	112	51.6	38	3	US-08-630-916A-31	Sequence 31, Appl
39	112	51.6	38	3	US-08-630-916A-32	Sequence 32, Appl
40	111	51.2	38	3	US-08-539-205A-2	Sequence 2, Appl
41	111	51.2	735	3	US-09-392-163A-2	Sequence 2, Appl
42	111	51.2	735	4	US-08-630-916A-24	Sequence 24, Appl
43	108	49.8	38	3	US-08-348-518C-15	Sequence 15, Appl
44	108	49.8	38	3	US-08-476-509B-15	Sequence 15, Appl
45	108	49.8	38	3	US-08-476-509B-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-895-601-6  
; Sequence 6, Application US/08895601  
; Patent No. 6060262  
; GENERAL INFORMATION:  
; APPLICANT: Beer-Romero, Peggy  
; APPLICANT: Strack, Peter J.  
; APPLICANT: Glass, Susan J.  
; APPLICANT: Rolfe, Mark  
; TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,  
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/895.601  
; FILING DATE: 16-JUL-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MIV-096.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 927 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-895-601-6

Query Match 96.3%; Score 209; DB 3; Length 927;  
Best Local Similarity 97.4%; Pred. No. 4.8e-21;  
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38  
Db 218 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 255

```
RESULT 2
US-08-630-916A-18
; Sequence 18, Application US/08630916A
; Patent No. 601137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-916A-18

Query Match      92.2%; Score 200; DB 3; Length 38;
Best Local Similarity 94.7%; Pred. No. 2.1e-21;
Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDL 38
Db      1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDL 38

RESULT 3
US-09-066-074-11
; Sequence 11, Application US/09066074
; Patent No. 5952467
; GENERAL INFORMATION:
; APPLICANT: Hunter et al., Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/555,912A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,074
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/555,912
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; CLONE: Nedd4/Mouse
US-09-066-074-11

Query Match      89.4%; Score 194; DB 2; Length 38;
Best Local Similarity 86.8%; Pred. No. 1.5e-20;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDL 38
Db      1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPDPDDL 38

RESULT 4
US-08-555-912A-11
; Sequence 11, Application US/08555912A
; Patent No. 5972697
; GENERAL INFORMATION:
; APPLICANT: Hunter et al., Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,912A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```







;;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-20

Query Match 67.3%; Score 146; DB 3; Length 38;  
Best Local Similarity 64.9%; Pred. No. 9.7e-14;  
Matches 24; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Oy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRTPDQN 37  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRTPDQN 37

RESULT 11  
US-08-348-518C-14  
; Sequence 14, Application US/08348518C  
; Patent No. 6022740  
; GENERAL INFORMATION:  
; APPLICANT: SUDOL, MARIUS  
; APPLICANT: PEER, BORK  
; APPLICANT: HENRY, CHEN  
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/348,518C  
; FILING DATE: 01-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 343-1684  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO

;;  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: Mouse-2  
; IMMEDIATE SOURCE:  
; CLONE: Nedd4  
US-08-348-518C-14

Query Match 67.3%; Score 146; DB 3; Length 38;  
Best Local Similarity 64.9%; Pred. No. 9.7e-14;  
Matches 24; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Oy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRTPDQN 37  
|||||:|||||:|||||:|||||:|||||:|||||:  
Db 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRTPDQN 37

RESULT 12  
US-08-476-509B-14  
; Sequence 14, Application US/08476509B  
; Patent No. 6034212  
; GENERAL INFORMATION:  
; APPLICANT: SUDOL, MARIUS  
; APPLICANT: PEER, BORK  
; APPLICANT: HENRY, CHEN  
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,509B  
; FILING DATE: 01-DEC-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: Mouse-2  
; IMMEDIATE SOURCE:  
; CLONE: Nedd4  
US-08-476-509B-14

Query Match 67.3%; Score 146; DB 3; Length 38;  
Best Local Similarity 64.9%; Pred. No. 9.7e-14;  
Matches 24; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Oy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRTPDQN 37  
|||||:|||||:|||||:|||||:|||||:|||||:

Db 1 SGLPPGWEEKQDRGRSYYVDHNSKTTTWSKPTMQDD 37

## RESULT 13

US-08-630-916A-21  
; Sequence 21, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-21

Query Match 65.0%; Score 141; DB 3; Length 38;  
Best Local Similarity 68.6%; Pred. No. 5e-13;  
Matches 24; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SGLPPGWEEKQDRGRSYYVDHNSKTTTWSKPTMQ 35

Db 1 SGLPPGWEEKQDRGRSYYVDHNSRTTWTKTIVQ 35

## RESULT 14

US-08-539-205A-6  
; Sequence 6, Application US/08539205A  
; Patent No. 6001619  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; APPLICANT: Caligiuri, Maureen  
; APPLICANT: Nefsky, Bradley  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/539,205A  
; FILING DATE: 04-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-539-205A-6

Query Match 65.0%; Score 141; DB 3; Length 834;  
Best Local Similarity 70.6%; Pred. No. 1.9e-11;  
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEEKODILGRTYYVNHESRTOWKRTPOD 36

Db 74 LPPGWEEKVDNLGRTYYVNHNRRTTOWHRRPSLMD 107

## RESULT 15

US-09-392-163A-6  
; Sequence 6, Application US/09392163A  
; Patent No. 6503742  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; APPLICANT: Caligiuri, Maureen  
; APPLICANT: Nefsky, Bradley  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/392,163A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/539,205  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-392-163A-6

Query Match 65.0%; Score 141; DB 4; Length 834;  
 Best Local Similarity 70.6%; Pred. No. 1.9e-11;  
 Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDQ 36  
 |||||:|||||:|||||:|||||:  
 Db 74 LPPGWEKVDNLGRYYVNHNNRTQWHRPSLMD 107

Search completed: October 13, 2005, 14:03:05  
 Job time : 19.2202 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compuquen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 81.011 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-4  
Perfect score: 217  
Sequence: 1 SGLPPGWEERQDILGRYYVNHESRRTQWKRPDQNL 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

May 20 1964

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA: \*

1:	/cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pcp:
2:	/cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pcp:
3:	/cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pcp:
4:	/cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pcp:
5:	/cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pcp:
6:	/cgn2_6/prodata/1/pubpaa/ACTU5_PUBCOMB.pcp:
7:	/cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pcp:
8:	/cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp:
9:	/cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pcp:
10:	/cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pcp:
11:	/cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pcp:
12:	/cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pcp:
13:	/cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pcp:
14:	/cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pcp:
15:	/cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pcp:
16:	/cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pcp:
17:	/cgn2_6/prodata/1/pubpaa/US10E_PUBCOMB.pcp:
18:	/cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pcp:
19:	/cgn2_6/prodata/1/pubpaa/US11A_PUBCOMB.pcp:
20:	/cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pcp:
21:	/cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pcp:
22:	/cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	217	100.0	38	14	US-10-307-956-4	Sequence 4, Appli
2	209	96.3	38	14	US-10-307-956-3	Sequence 3, Appli
3	209	96.3	733	14	US-10-097-534-16	Sequence 16, Appl
4	209	96.3	927	14	US-10-097-534-15	Sequence 15, Appl
5	200	92.2	38	14	US-10-185-050-18	Sequence 18, Appl
6	194	89.4	38	14	US-10-256-865-36	Sequence 36, Appl
7	194	89.4	38	14	US-10-185-050-17	Sequence 17, Appl
8	194	89.4	38	16	US-10-687-361-11	Sequence 11, Appl
9	194	89.4	38	17	US-10-716-379-11	Sequence 11, Appl
10	194	89.4	38	17	US-10-616-410-11	Sequence 11, Appl
11	194	89.4	38	17	US-10-648-631-11	Sequence 11, Appl

12	146	67.3	38	14	US-10-185-050-20	Sequence 20, Appl
13	141	65.0	38	14	US-10-185-050-21	Sequence 21, Appl
14	141	65.0	834	14	US-10-313-955-6	Sequence 6, Appl
15	141	65.0	854	14	US-10-205-823-277	Sequence 277, App
16	141	65.0	854	14	US-10-287-218-3	Sequence 3, Appl
17	141	65.0	854	16	US-10-474-291-3	Sequence 3, Appl
18	141	65.0	854	20	US-11-051-454-277	Sequence 277, App
19	141	65.0	911	14	US-10-205-823-279	Sequence 279, App
20	141	65.0	911	20	US-11-051-454-279	Sequence 279, App
21	141	65.0	923	18	US-10-450-763-31439	Sequence 31439, A
22	141	65.0	975	16	US-10-779-271-3	Sequence 3, Appl
23	141	65.0	995	14	US-10-097-534-9	Sequence 9, Appl
24	141	65.0	995	14	US-10-205-823-275	Sequence 275, App
25	141	65.0	995	20	US-11-051-454-275	Sequence 275, App
26	141	65.0	1071	18	US-10-450-763-42131	Sequence 42131, A
27	138	63.6	1082	20	US-11-097-143-16857	Sequence 16857, A
28	135	62.2	38	14	US-10-185-050-36	Sequence 36, Appl
29	133	61.3	38	14	US-10-185-050-132	Sequence 132, App
30	133	61.3	38	17	US-10-785-819-168	Sequence 168, App
31	133	61.3	725	14	US-10-185-050-126	Sequence 126, App
32	132	60.8	38	14	US-10-256-865-37	Sequence 37, Appl
33	132	60.8	38	14	US-10-185-050-19	Sequence 19, Appl
34	132	60.8	38	14	US-10-307-956-7	Sequence 7, Appl
35	132	60.8	38	16	US-10-687-361-12	Sequence 12, Appl
36	132	60.8	38	17	US-10-716-379-12	Sequence 12, Appl
37	132	60.8	38	17	US-10-616-410-12	Sequence 12, Appl
38	132	60.8	38	17	US-10-648-631-12	Sequence 12, Appl
39	126	58.1	832	14	US-10-032-585-7296	Sequence 7296, Ap
40	123	56.7	759	14	US-10-128-714-3162	Sequence 3162, Ap
41	123	56.7	766	14	US-10-313-955-4	Sequence 4, Appl
42	123	56.7	869	14	US-10-128-714-8162	Sequence 8162, Ap
43	122	56.2	38	14	US-10-307-956-26	Sequence 26, Appl
44	122	56.2	683	14	US-10-185-050-46	Sequence 46, Appl
45	122	56.2	684	14	US-10-097-534-11	Sequence 11, Appl

## ALIGNMENTS

```

RESULT 1
US-10-307-956-4
; Sequence 4, Application US/10/07956
; Publication NO. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoeckstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCES: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307.956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-4

```

```
Query Match      100.0%; Score 217; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 SGLPGWEERQDILGRYYVNHESRPTQWKRPDQNL 38  
|||||

Db 1 SGLPGWEERQDILGRYYVNHESRPTQWKRPDQNL 38  
|||||

## RESULT 2

US-10-307-956-3  
; Sequence 3, Application US/10307956  
; Publication No. US2003011907A1  
; GENERAL INFORMATION:

; APPLICANT: Hoekstra, Merl F.

; APPLICANT: Xie, Weilin

; APPLICANT: Murray, Brian

; APPLICANT: Mercurio, Frank

; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL

; FILE REFERENCE: 860098.433

; CURRENT APPLICATION NUMBER: US/10/307,956

; PRIOR FILING DATE: 2002-12-02

; PRIOR APPLICATION NUMBER: US/09/385,918

; PRIOR FILING DATE: 1999-08-30

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 38

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-307-956-3

Query Match 96.3%; Score 209; DB 14; Length 38;

Best Local Similarity 97.4%; Pred. No. 2.1e-20;

Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

RESULT 3

US-10-097-534-16

; Sequence 16, Application US/10097534

; Publication No. US20030049607A1

; GENERAL INFORMATION:

; APPLICANT: GREENER, TSVIKA

; APPLICANT: MOSKOWITZ, HAIM

; APPLICANT: REISS, YUVAL

; APPLICANT: ALROY, IRIS

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL

; FILE REFERENCE: PLV-001.01

; CURRENT APPLICATION NUMBER: US/10/097,534

; PRIOR FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/275,224

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: 60/308,958

; PRIOR FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: 60/340,170

; PRIOR FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 16

; LENGTH: 733

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-097-534-16

Query Match 96.3%; Score 209; DB 14; Length 733;

Best Local Similarity 97.4%; Pred. No. 4.7e-19;

Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

Db 191 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 228

RESULT 4

US-10-097-534-15

; Sequence 15, Application US/10097534

; Publication No. US20030049607A1

; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/340,170  
; PRIOR FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 927  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-534-15

Query Match 96.3%; Score 209; DB 14; Length 927;

Best Local Similarity 97.4%; Pred. No. 6e-19;

Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 38

Db 218 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDQNL 255

RESULT 5

US-10-185-050-18

; Sequence 18, Application US/10185050

; Publication No. US2003007577A1

; GENERAL INFORMATION:

; APPLICANT: Pirozzi, Gregorio

; Kay, Brian K.

; Fowlkes, Dana M.

; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

; NUMBER OF SEQUENCES: 233

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/185,050

; FILING DATE: 28-Jun-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/826,516

; FILING DATE: 03-Apr-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: MISROCK, S. LESLIE

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-208-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 896-8864/9741

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-185-050-18

Query Match          92.2%; Score 200; DB 14; Length 38;
Best Local Similarity 94.7%; Pred. No. 3.4e-19;
Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-10-256-865-36
; Sequence 36, Application US/10256865
; Publication No. US20030049672A1
; GENERAL INFORMATION:
; APPLICANT: Xiao Zhen Zhou
; TITLE OF INVENTION: Methods and Compositions for Regulating
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 1440.1034-000
; CURRENT APPLICATION NUMBER: US/10/256.865
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: PRIOR APPLICATION NUMBER: US/09/252,404
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 38
; TYPE: PPT
; ORGANISM: mouse
US-10-256-865-36

Query Match          89.4%; Score 194; DB 14; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.1e-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPSPDDDL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-10-185-050-17
; Sequence 17, Application US/10185050
; Publication No. US2003007577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
```

```
;
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-185-050-17

Query Match          89.4%; Score 194; DB 14; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.1e-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPSPDDDL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-10-687-361-11
; Sequence 11, Application US/10687361
; Publication No. US20040101896A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Lu, Kun Ping
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-078
; CURRENT APPLICATION NUMBER: US/10/687,361
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PPT
; ORGANISM: Mus musculus
US-10-687-361-11

Query Match          89.4%; Score 194; DB 16; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.1e-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDNL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 SPLPPGWEERQDVLGRYYVNHESRRRTQWKRPSPDDDL 38
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-10-716-379-11
; Sequence 11, Application US/10716379
; Publication No. US20050027107A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; APPLICANT: Kun Ping, Lu
; APPLICANT: Hanes, Steven D.
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-085
; CURRENT APPLICATION NUMBER: US/10/716,379
```

```

; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-716-379-11

Query Match      89.4%; Score 194; DB 17; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.le-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY   1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQNL 38
    |||||||:|||||||:|||||||:|||||||:|:|
Db   1 SPLPGWEERQDVLRYYVNHESRRRTQWKRPSPDDDL 38
    |||||||:|||||||:|||||||:|||||||:|:|

RESULT 10
US-10-616-410-11
; Sequence 11, Application US/10616410
; Publication No. US20050033032A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-043
; CURRENT APPLICATION NUMBER: US/10/616,410
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US 09/275,900
; PRIOR FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-616-410-11

Query Match      89.4%; Score 194; DB 17; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.le-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY   1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQNL 38
    |||||||:|||||||:|||||||:|||||||:|:|
Db   1 SPLPGWEERQDVLRYYVNHESRRRTQWKRPSPDDDL 38
    |||||||:|||||||:|||||||:|||||||:~

RESULT 11
US-10-648-631-11
; Sequence 11, Application US/10648631
; Publication No. US20050049404A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, Tony
; TITLE OF INVENTION: NIMA INTERACTING PROTEINS
; FILE REFERENCE: 66671-044
; CURRENT APPLICATION NUMBER: US/10/648,631
; CURRENT FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: US 10/616,410
; PRIOR FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-648-631-11

Query Match      89.4%; Score 194; DB 17; Length 38;
Best Local Similarity 86.8%; Pred. No. 2.le-18;
Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY   1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQNL 38
    |||||||:|||||||:|||||||:|||||||:|:|
Db   1 SPLPGWEERQDVLRYYVNHESRRRTQWKRPSPDDDL 38
    |||||||:|||||||:|||||||:|||||||:~

RESULT 12
US-10-185-050-20
; Sequence 20, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio Kay, Brian K. Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-185-050-20

Query Match      67.3%; Score 146; DB 14; Length 38;
Best Local Similarity 64.9%; Pred. No. 5.5e-12;
Matches 24; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY   1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQN 37
    |||||||:|||||||:|||||||:|||||||:~
Db   1 SGLPPGWEERQDRLGRSYYVDHNSKTTTWSKPTMQDD 37
    |||||||:|||||||:|||||||:|||||||:~

RESULT 13
US-10-185-050-21
; Sequence 21, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio Kay, Brian K. Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

```



## POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

```
;
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-185-050-21

Query Match 65.0%; Score 141; DB 14; Length 38;
Best Local Similarity 68.6%; Pred. No. 2.5e-11;
Matches 24; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SGLPPGWEERQDILGRTYYVNHESRRTQWKRPPTQ 35
| | | | | | | | | | | | | | | | | | | |
Db 1 SGLPPGWEERQDILGRTYYVNHESRRTQWKRPPTQ 35
| | | | | | | | | | | | | | | | | | | |

RESULT 14
US-10-313-955-6
; Sequence 6, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Caligiuri, Maureen
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313,955
; FILING DATE: 05-Dec-2002
;
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-313-955-6

Query Match 65.0%; Score 141; DB 14; Length 834;
Best Local Similarity 70.6%; Pred. No. 6.5e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRTYYVNHESRRTQWKRPPTQ 36
| | | | | | | | | | | | | | | | | | | |
Db 74 LPPGWEERQDILGRTYYVNHESRRTQWKRPPTQ 107
| | | | | | | | | | | | | | | | | | | |

RESULT 15
US-10-205-823-277
; Sequence 277, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-277

Query Match 65.0%; Score 141; DB 14; Length 854;
Best Local Similarity 70.6%; Pred. No. 6.7e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

Qy 3 LPPGWEERODILGRYYVNHESRRTOWKRTPOD 36  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 74 LPPGWEKVDNLGRYYVNHNNRTTQWHRPSLMD 107

Search completed: October 13, 2005, 14:50:00  
Job time : 81.011 secs





A;Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.160  
A;Experimental source: BAC clone B24B19; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B24B19.160  
A;Map position: 6  
A;Introns: 11/1; 24/1; 59/2; 110/1; 783/2  
C;Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology  
F;235-276/Domain: WW repeat homology <WW1>  
F;334-371/Domain: WW repeat homology <WW2>  
F;393-430/Domain: WW repeat homology <WW3>

Query Match 59.9%; Score 130; DB 2; Length 815;  
Best Local Similarity 71.0%; Pred. No. 1.2e-08;  
Matches 22; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 33  
||| : ||||| : |||  
Db 241 LPAGWERREDHLGRYYVDHNSRTTSMNRPT 271

RESULT 7  
S66562  
ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)  
N;Alternate names: E6-AP-like protein ubiquitin ligase  
C;Species: Schizosaccharomyces pombe  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S66562; T45159  
R;Netsky, B.; Beach, D.  
EMBO J. 15, 1301-1312, 1996  
A;Title: Pub1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25  
A;Reference number: S66562; MUID:96205868; PMID:8635463  
A;Accession: S66562  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-766 <NEF>  
A;Cross-references: UNIPROT:Q92462; GB:U66716; NID:G1519443; PIDN:AAB07514.1; PID:G15194  
R;Netsky, B.S.; Beach, D.  
submitted to the EMBL Data Library, August 1996  
A;Description: Pub1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of  
A;Reference number: Z22935  
A;Accession: T45159  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-766 <NE2>  
A;Cross-references: EMBL:Y07592; PIDN:CAA68867.1  
C;Genetics:  
A;Gene: pub1  
C;Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology  
C;Keywords: ligase  
F;205-242/Domain: WW repeat homology <WW1>  
F;288-325/Domain: WW repeat homology <WW2>  
F;345-382/Domain: WW repeat homology <WW3>

Query Match 56.7%; Score 123; DB 1; Length 766;  
Best Local Similarity 70.0%; Pred. No. 9.1e-08;  
Matches 21; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 32  
||| : ||||| : |||  
Db 207 LPPGWERTDNLGRYYVDHNRSTTWIRP 236

RESULT 8  
T37545  
ubiquitin-protein ligase (EC 6.3.2.19) pub1 [validated] - fission yeast (Schizosaccharom  
N;Alternate names: ubiquitin ligase Pub1  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T37545; T48655  
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A;Reference number: Z21722  
A;Accession: T37545

A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-767 <MCL>  
A;Cross-references: UNIPROT:Q92462; EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:SF  
A;Experimental source: strain 972h-; cosmid c11G7  
R;Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.  
Mol. Gen. Genet. 254, 520-538, 1997  
A;Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pub1 ut  
A;Reference number: 207985; MUID:97340937; PMID:9197411  
A;Accession: T48655  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-767 <SAL>  
A;Cross-references: EMBL:U62795; NID:G2262192; PIDN:AAB63350.1; PID:G2262193  
A;Experimental source: strain J227  
C;Genetics:  
A;Gene: pub1; SPDB:SPAC11G7.02  
A;Map position: 1  
A;Introns: 6/2; 14/1; 62/2  
C;Function: <UBI>  
A;Description: EC 6.3.2.19 [validated, MUID:96205868]  
C;Function: <CYC>  
C;Function: <TOL>  
A;Description: required for low pH-tolerance [validated, MUID:97340937]  
C;Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology  
C;Keywords: cell cycle control; ligase  
F;205-242/Domain: WW repeat homology <WW1>  
F;288-325/Domain: WW repeat homology <WW2>  
F;345-382/Domain: WW repeat homology <WW3>

Query Match 56.7%; Score 123; DB 2; Length 767;  
Best Local Similarity 70.0%; Pred. No. 9.1e-08;  
Matches 21; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 32  
||| : ||||| : |||  
Db 207 LPPGWERTDNLGRYYVDHNRSTTWIRP 236

RESULT 9  
T47801  
hypothetical protein F24G16.40 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T47801  
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.,  
submitted to the Protein Sequence Database, February 2000  
A;Reference number: 224477  
A;Accession: T47801  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1616 <DAN>  
A;Cross-references: UNIPROT:Q9M1Z7; EMBL:AL138647  
A;Experimental source: cultivar Columbia; BAC clone F24G16  
C;Genetics:  
A;Map position: 3  
A;Introns: 7/1; 205/3; 250/1; 656/3; 713/3; 794/3; 964/3; 1232/3; 1298/3; 1445/3; 1523/3  
A;Note: F24G16.40  
F;508-545/Domain: WW repeat homology <WWR>

Query Match 49.5%; Score 107.5; DB 2; Length 1616;  
Best Local Similarity 48.5%; Pred. No. 1.9e-05;  
Matches 16; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 34  
||| : ||||| : |||  
Db 510 LPPGWEKRADAVTGKSYIIDHTTKTTTWSHPCP 542

RESULT 10  
S68520  
Finl protein - human

C;Species: Homo sapiens (man)  
C;Date: 15-Nov-1996 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
C;Accession: S68520  
R;Lu, K.P.; Hanes, S.D.; Hunter, T.  
Nature 380, 544-547, 1996  
A;Title: A human peptidyl-prolyl isomerase essential for regulation of mitosis.  
A;Reference number: S68520; MUID:96195064; PMID:8606777  
A;Accession: S68520  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-163 <LUK>  
A;Cross-references: UNIPROT:Q13526; EMBL:U49070; NID:U49070; PIDN:AAC50492.1; PID:g133  
C;Genetics:  
A;Gene: GDB:PIN1; dod  
A;Cross-references: GDB:5218381  
C;Superfamily: Yeast ESS1 protein; WW repeat homology  
F;5-43/Domain: WW repeat homology <WW1>  
  
Query Match 44.9%; Score 97.5; DB 2; Length 163;  
Best Local Similarity 53.1%; Pred. No. 3.2e-05;  
Matches 17; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
  
Qy 3 LPPGWEERQD-ILGRYYVNHESRRRTQWKRP 33  
|||||:| |||||:| :|||:  
Db 7 LPPGWEKRMSRSGRVYFNHITNASQWERPS 38  
|||||:| |||||:| :|||:  
  
RESULT 11  
JC7136  
peptidylprolyl isomerase (EC 5.2.1.8) - mouse  
N;Alternate names: parvulins, Pin1 protein  
C;Species: Mus musculus (house mouse)  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: JC7136  
R;Fujimori, F.; Takahashi, K.; Uchida, C.; Uchida, T.  
Biochem. Biophys. Res. Commun. 265, 658-663, 1999  
A;Title: Mice lacking Pin1 develop normally, but are defective in entering cell cycle fr  
A;Reference number: JC7136; MUID:20070807; PMID:10600477  
A;Accession: JC7136  
A;Molecule type: mRNA  
A;Residues: 1-165 <FUJ>  
A;Cross-references: UNIPROT:Q9QUR7; DBJ:AB009691; NID:g6468199; PIDN:BAAR7037.1; PID:96  
C;Comment: This protein is a member of peptidyl prolyl cis/trans isomerase family. It is  
tes in oligopeptides, controls G2/M transition of HeLa cells, regulates entry into G1 an  
C;Genetics:  
A;Gene: Pin1  
C;Superfamily: Yeast ESS1 protein; WW repeat homology  
C;Keywords: cis-trans-isomerase; mitosis  
F;5-43/Domain: WW repeat homology <WWR>  
  
Query Match 44.9%; Score 97.5; DB 2; Length 165;  
Best Local Similarity 53.1%; Pred. No. 3.2e-05;  
Matches 17; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
  
Qy 3 LPPGWEERQD-ILGRYYVNHESRRRTQWKRP 33  
|||||:| |||||:| :|||:  
Db 7 LPPGWEKRMSRSGRVYFNHITNASQWERPS 38  
|||||:| |||||:| :|||:  
  
RESULT 12  
JC7507  
45K WW domain-containing protein - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C;Accession: JC7507  
R;Valverde, P.  
Biochem. Biophys. Res. Commun. 276, 990-998, 2000  
A;Title: Cloning, expression, and mapping of hmw45, a novel human WW domain-containing  
A;Reference number: JC7507; MUID:20483619; PMID:11027580  
A;Accession: JC7507  
A;Molecule type: mRNA  
A;Residues: 1-383 <VAL>  
A;Cross-references: GB:AJ292969

C;Comment: This protein participates in a variety of cellular processes, such as protein  
C;Genetics:  
A;Gene: ww45  
A;Map position: 14q13-21, 14q21-23  
C;Keywords: coiled coil; myristylation; protein degradation; transcription  
  
Query Match 41.9%; Score 91; DB 2; Length 383;  
Best Local Similarity 38.9%; Pred. No. 0.00055;  
Matches 14; Conservative 6; Mismatches 16; Indels 0; Gaps 0;  
  
Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 38  
|||||:| |||||:| :|||:  
Db 201 LPPGWSYDWTMRGRKYYIDHTNTTTHWSHPLERGL 236  
|||||:| |||||:| :|||:  
  
RESULT 13  
JC7508  
45K WW domain-containing protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C;Accession: JC7508  
R;Valverde, P.  
Biochem. Biophys. Res. Commun. 276, 990-998, 2000  
A;Title: Cloning, expression, and mapping of hmw45, a novel human WW domain-containing  
A;Reference number: JC7507; MUID:20483619; PMID:11027580  
A;Accession: JC7508  
A;Molecule type: mRNA  
A;Residues: 1-386 <VAL>  
A;Cross-references: UNIPROT:QVVEB2; GB:AJ292968  
C;Comment: This protein participates in a variety of cellular processes, such as protein  
during early developmental stage.  
C;Genetics:  
A;Gene: ww45  
C;Keywords: coiled coil; myristylation; protein degradation; transcription  
  
Query Match 41.9%; Score 91; DB 2; Length 386;  
Best Local Similarity 38.9%; Pred. No. 0.00055;  
Matches 14; Conservative 6; Mismatches 16; Indels 0; Gaps 0;  
  
Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 38  
|||||:| |||||:| :|||:  
Db 202 LPPGWSYDWTMRGRKYYIDHTNTTTHWSHPLERGL 237  
|||||:| |||||:| :|||:  
  
RESULT 14  
T14152  
synaptic scaffolding protein S-SCAM - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 15-Mar-2004  
C;Accession: T14152  
R;Hirao, K.; Hata, Y.; Ide, N.; Takeuchi, M.; Irie, M.; Yao, I.; Deguchi, M.; Toyoda, A.  
J. Biol. Chem. 273, 21105-21110, 1998  
A;Title: A novel multiple PDZ domain-containing molecule interacting with N-methyl-d-aspa  
A;Reference number: Z17889; MUID:98361985; PMID:9694864  
A;Accession: T14152  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1277 <HIR>  
A;Cross-references: EMBL:AF034863; NID:g3411052; PID:g3411053; PIDN:AAC31124.1  
C;Genetics:  
A;Gene: S-SCAM  
C;Function:  
A;Description: may assemble receptors and cell adhesion proteins at synaptic junctions  
F;302-339/Domain: WW repeat homology <WW1>  
F;348-385/Domain: WW repeat homology <WW2>  
  
Query Match 40.1%; Score 87; DB 2; Length 1277;  
Best Local Similarity 53.3%; Pred. No. 0.0064;  
Matches 16; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
  
Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRP 32  
|||||:| |||||:| :|||:  
Db 350 LPYGWEKIDDPYGYVVDHINRRTQFNP 379  
|||||:| |||||:| :|||:



**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 72.381 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-4

Perfect score: 217

Sequence: 1 SGLPFGWEQDILGRYYVNHESRRQWKRPQDNL 38

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209	96.3	1000	1 NED4_HUMAN	P46934 homo sapien
2	204	94.0	455	2 Q9N134	Q9n134 oryctolagus
3	194	89.4	887	1 NED4_MOUSE	P46935 mus musculus
4	190	87.6	887	1 NED4_RAT	Q62940 rattus norv
5	141	65.0	442	2 Q8OU03	Q8ou03 mus musculus
6	141	65.0	820	2 Q9NT88	Q9nt88 homo sapien
7	141	65.0	835	2 Q8BT19	Q8btr9 mus musculus
8	141	65.0	854	2 Q9H2W4	Q9h2w4 homo sapien
9	141	65.0	855	2 Q9BW58	Q9bw58 homo sapien
10	141	65.0	855	2 Q99PK2	Q99pk2 mus musculus
11	141	65.0	855	2 Q8CF10	Q8cf10 mus musculus
12	141	65.0	875	2 Q641N8	Q641n8 mus musculus
13	141	65.0	911	2 Q8NSA7	Q8nsa7 homo sapien
14	141	65.0	947	2 Q7Z5F1	Q7z5f1 homo sapien
15	141	65.0	955	2 Q96PU5	Q96pu5 homo sapien
16	141	65.0	967	2 Q7Z5F2	Q7z5f2 homo sapien
17	141	65.0	975	2 Q7Z5N3	Q7z5n3 homo sapien
18	141	65.0	995	2 Q43165	Q43165 homo sapien
19	139	64.1	917	2 Q7QE76	Q7qe76 anopheles g
20	138	63.6	571	2 Q8QGJ2	Q8qgj2 gallus gall
21	138	63.6	793	2 Q7KUR2	Q7kur2 drosophila
22	138	63.6	834	2 Q9SR64	Q9sr64 drosophila
23	138	63.6	938	2 Q85TQ0	Q8stq0 drosophila
24	138	63.6	956	2 Q81QR6	Q81qr6 drosophila
25	138	63.6	1007	2 Q9VVI3	Q9vvi3 drosophila
26	136	62.7	970	2 Q6DIR6	Q6dir6 xenopus tro
27	135	62.2	971	2 Q42573	Q42573 xenopus lae
28	135	62.2	971	2 Q6GMD5	Q6gmd5 xenopus lae
29	135	62.2	1048	2 Q7PQR5	Q7pqr5 anopheles g
30	132	60.8	809	1 RSP5	P39940 saccharomyc
31	132	60.8	817	2 Q75AI2	Q75ai2 ashbya goss

#### RESULT 1

NED4\_HUMAN  
ID NED4\_HUMAN STANDARD; PRT; 1000 AA.  
AC P46934;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).  
GN Name=NED4; Synonyms=KIAA0093;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP CONCEPTUAL TRANSLATION OF 1-73.  
RA Hinz U.;  
RL Unpublished observations (JUN-2003).  
RN [2]  
RP SEQUENCE OF 74-1000 FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=95308325; PubMed=7788527;  
RA Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,  
RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. III.  
RT The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by  
RT analysis of cDNA clones from human cell line KG-1.";  
RNA Res. 2:37-43(1995).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
CC then directly transfers the ubiquitin to targeted substrates (By  
CC similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Interacts with UBE2D2. Binds SCNN1A, SCNN1B and SCNN1G  
CC (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
CC thiolester formation.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -!- SIMILARITY: Contains 4 WW domains.  
CC -!- CAUTION: The sequence of the N-terminus was deduced from the  
CC genomic sequence.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AC039057; -, NOT\_ANNOTATED\_CDS.  
CC EMBL; D42055; BAA07655.1; -.

32 132 60.8 819 2 Q6CNC7 Q6cnc7 kluyveromyc  
33 132 60.8 822 2 Q6FN71 Q6fn71 candida gla  
34 131 60.4 781 2 Q6BT41 Q6bt41 debaryomyce  
35 130 59.9 786 1 PUB3\_SCHPO O14326 schizosacch  
36 130 59.9 787 2 Q6M906 Q6m906 neurospora  
37 130 59.9 806 2 Q7RV01 Q7rv01 neurospora  
38 123 56.7 767 1 PUB1\_SCHPO Q92462 schizosacch  
39 122 56.2 918 1 WWP1\_MOUSE Q8bzz3 mus musculu  
40 122 56.2 922 1 WWP1\_HUMAN Q9H0M0 homo sapien  
41 121.5 56.0 870 1 WWP2\_HUMAN O00308 homo sapien  
42 121.5 56.0 870 1 WWP2\_MOUSE Q9dbh0 mus musculu  
43 121 55.8 806 2 Q6P289 Q6p289 mus musculu  
44 121 55.8 854 2 Q6C5H0 Q6c5h0 yarrowia li  
45 121 55.8 864 1 ITCH\_MOUSE Q8c863 mus musculu

#### ALIGNMENTS

```

DR HSSP; Q62940; 115H.
DR Genew; HGNC:7727; NEDD4.
DR H-invDB; HIX0012269; -.
DR MIM; 602278; -.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2 CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTc; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2 DOMAIN 1; 1.
DR PROSITE; PS00004; C2 DOMAIN 2; 1.
DR PROSITE; PS50004; C2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS00020; WW DOMAIN 2; 4.
KW Ligase; Repeat; Ubl conjugation; Ubl conjugation pathway.
FT DOMAIN 9 44 Pro-rich.
FT DOMAIN 77 81 Poly-Ser.
FT DOMAIN 106 207 C2 domain.
FT DOMAIN 285 288 Poly-Gln.
FT DOMAIN 291 324 WW 1.
FT DOMAIN 448 481 WW 2.
FT DOMAIN 521 554 WW 3.
FT DOMAIN 573 606 WW 4.
FT DOMAIN 665 1000 HECT.
FT BINDING 967 967 Ubiquitin (By similarity).
SQ SEQUENCE 1000 AA; 114936 MW; 372808BE50C149CB CRC64;

Query Match 96.3%; Score 209; DB 1; Length 1000;
Best Local Similarity 97.4%; Pred. No. 8.1e-19;
Matches 37; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPPTQDNL 38
DB 291 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPPTQDNL 328

RESULT 2
O9N134 PRELIMINARY; PRT; 455 AA.
ID Q9N134
AC Q9N134;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-protein ligase Nedd4 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21113801; PubMed=11181416;
RA Velazquez H., Silva T., Andujar E., Desir G.V., Ellison D.H.,
RA Greger R.;
RT "The distal convoluted tubule of rabbit kidney does not express a
RT functional sodium channel."
RL Am. J. Physiol. Renal Physiol. 280:F530-F539(2001).
DR EMBL; AF229024; AAF45194.1; -.
DR HSSP; Q62940; 115H.
DR GO; GO:0016874; F:ligase activity; IEA.
DR InterPro; IPR008973; C2 CaLB.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 4.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS50020; WW DOMAIN 2; 4.
KW Ligase.
FT NON_TER 1 1

```

---

```

FT NON TER 455 455
SQ SEQUENCE 455 AA; 51811 MW; D1C4549B5E22298 CRC64;

Query Match 94.0%; Score 204; DB 2; Length 455;
Best Local Similarity 94.7%; Pred. No. 1.6e-16;
Matches 36; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPPTQDNL 38
DB 98 SPLPPGWEERQDILGRYYVNHESRRRTQWKRPPTQDNL 135

RESULT 3
NED4_MOUSE STANDARD; PRT; 887 AA.
ID NED4_MOUSE
AC P46935; O08758; Q85GB3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN Name=Nedd4; Synonyms=Kiaa0093, Nedd-4, Nedd4a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=92328780; PubMed=1378265;
RA Kumar S., Tomooka Y., Noda M.;
RT "Identification of a set of genes with developmentally down-regulated
RT expression in the mouse brain."
RL Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
RN [2]
RP REVISIONS.
RA Kumar S.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.; SUBCELLULAR LOCATION, AND INTERACTION WITH UBE2D2.
RX STRAIN=C57BL/6 X CBA;
RX MEDLINE=9736076; PubMed=9182527; DOI=10.1074/jbc.272.24.15085;
RA Hatakeyama S., Jensen J.P., Weissman A.M.;
RT "Subcellular localization and ubiquitin-conjugating enzyme (E2)
RT interactions of mammalian HECT family ubiquitin protein ligases."
RL J. Biol. Chem. 272:15085-15092(1997).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tonaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gutincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Resole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Oi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Shibata K., Shinagawa A.,

```

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22579291; PubMed=12693553;  
 RA Okazaki N., Kikuno R., Inamoto S., Aizawa H., Yuasa S.,  
 RA Nakajima D., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 10:35-48(2003).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates (By  
 CC similarity).  
 CC -!- FUNCTION: Involved in the embryonic development and  
 CC differentiation of the central nervous system.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B and SCNN1G (By similarity).  
 CC Interacts with UBE2D2.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Brain.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: Contains 3 WW domains.  
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to a  
 CC frameshift in position 12.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D85414; BAA12803.1; -;  
 DR EMBL; U96635; AAB63360.1; ALT FRAME.  
 DR EMBL; AK088620; BAC40458.1; -;  
 DR EMBL; AK088767; BAC40558.1; -;  
 DR EMBL; AK122203; BAC65485.1; -;  
 DR HSSP; Q62940; 115H.  
 DR MGD; MGI:97297; Nedd4.  
 DR GO; GO:0005829; C:cytosol; IDA.  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IPI.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2\_CaLB.  
 DR InterPro; IPR000569; HECT.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 3.  
 DR PRINTS; PR00360; C2DOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS50237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 3.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 3.  
 KW Ligase; Repeat; Ubl conjugation; Ubl conjugation pathway.  
 FT DOMAIN 65 166 C2 domain.

FT DOMAIN 249 282 WW 1.  
 FT DOMAIN 405 438 WW 2.  
 FT DOMAIN 460 493 WW 3.  
 FT DOMAIN 552 887 HECT.  
 FT BINDING 854 854 Ubiquitin (By similarity).  
 SQ SEQUENCE 887 AA; 102705 MW; AE7DD3ED63986C50 CRC64;  
 Query Match 89.4%; Score 194; DB 1; Length 887;  
 Best Local Similarity 86.8%; Pred. No. 7e-17;  
 Matches 33; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 SGLPGWEEERODILGRYYVNHESRRRTQWKPTQDNL 38  
 Db 249 SPLPGWEEERQDVLGRYYVNHESRRRTQWKRPSPDDL 286  
 RESULT 4  
 NED4\_RAT  
 ID\_NED4\_RAT STANDARD; PRT; 887 AA.  
 AC Q62940;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE E3 ubiquitin-protein ligase Nedd4 (EC 6.3.2.-).  
 GN Name=Nedd4;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.  
 RC TISSUE=Lung;  
 RX MEDLINE=96221297; PubMed=8665844;  
 RA Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,  
 RA Rotin D.;  
 RT "WW domains of Nedd4 bind to the proline-rich PY motifs in the  
 RT epithelial Na+ channel deleted in Liddle's syndrome.";  
 RL EMBL J. 15:2371-2380(1996).  
 RN [2]  
 RP STRUCTURE BY NMR OF 452-499 IN COMPLEX WITH SCNN1B, AND INTERACTION  
 RP WITH SCNN1A; SCNN1B AND SCNN1G.  
 RX MEDLINE=21223577; PubMed=11323714; DOI=10.1038/87562;  
 RA Kanelis V., Rotin D., Forman-Kay J.D.;  
 RT "Solution structure of a Nedd4 WW domain-ENaC peptide complex.";  
 RL Nat. Struct. Biol. 8:407-412(2001).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates (By  
 CC similarity).  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Interacts with UBE2D2 (By similarity). Binds SCNN1A,  
 CC SCNN1B and SCNN1G.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Expression is highest in lung,  
 CC kidney and brain.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: Contains 3 WW domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U50842; AAB48949.1; -;  
 DR PIR; S70642; S70642.  
 DR PDB; 115H; NMR; W=450-499.

```
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS00020; WW_DOMAIN_2; 3.
KW 3D-structure; Ligase; Repeat; Ubl conjugation;
KW Ubl conjugation pathway.
FT DOMAIN 62 163 C2 domain.
FT DOMAIN 246 279 WW 1.
FT DOMAIN 402 435 WW 2.
FT DOMAIN 459 492 WW 3.
FT DOMAIN 551 887 HECT.
FT BINDING 854 854 Ubiqutin (By similarity).
FT TURN 463 464
FT STRAND 465 469
FT TURN 471 472
FT STRAND 475 479
FT TURN 480 483
FT STRAND 484 486
FT TURN 490 492
SQ SEQUENCE 887 AA; 102394 MW; D74BL097688CD9A1 CRC64;

Query Match 87.6%; Score 190; DB 1; Length 887;
Best Local Similarity 84.2%; Pred. No. 2.4e-16;
Matches 32; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLPPGWEERQDILGRYYVNHESRRRTQWKRPDPQDL 38
DB 246 SPLPGWEERQDVLGRYYVNHESRRRTQWKRPSPEDDL 283

RESULT 5
ID Q80U03 PRELIMINARY; PRT; 442 AA.
AC Q80U03; 2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MKTAA0439 protein (Fragment).
GN Name=MKIAA0439;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
CC -!- SIMILARITY: Contains 1 C2 domain.
DR EMBL; AK122283; BAC65565.1; -.
DR HSP; O13526; 1PIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
```

```
DR Pfam; PF00397; WW; 1.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 1.
FT NON TER 1
SQ SEQUENCE 442 AA; 50752 MW; DFF65B3BA6C2D21C CRC64;

Query Match 65.0%; Score 141; DB 2; Length 442;
Best Local Similarity 70.6%; Pred. No. 3.6e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDP 36
DB 173 LPPGWEKVDNLGRYYVNHNRSTQWHRPSLMD 206

RESULT 6
ID Q9NT88 PRELIMINARY; PRT; 820 AA.
AC Q9NT88; 2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKF2p434P2422 (Fragment).
GN Name=DKF2p434P2422;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 C2 domain.
DR EMBL; AL137469; CAB70754.1; -.
DR PIR; T46412; T46412.
DR HSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS00020; WW_DOMAIN_2; 3.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 820 AA; 95282 MW; 0FDB34B29B3F4123 CRC64;

Query Match 65.0%; Score 141; DB 2; Length 820;
Best Local Similarity 70.6%; Pred. No. 7.1e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDP 36
DB 144 LPPGWEKVDNLGRYYVNHNRSTQWHRPSLMD 177
```

```

RESULT 7
ID QBRT9 PRELIMINARY; PRT; 835 AA.
AC QBRT9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:AV3001F13 product:neural cell expressed,
DE developmentally down-regulated gene 4b, full insert sequence.
GN Name=Nedd4l;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazana M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1157-1171(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nihihi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,

```

```

RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR HSP; Q62940; 115H.
DR MGD; MGI:1933754; Nedd4l.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PRO0403; WNDOMAIN.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS00020; WW DOMAIN 2; 4.
SQ SEQUENCE 835 AA; 96410 MW; 81C7DD62722DDA9B CRC64;

Query Match 65.0%; Score 141; DB 2; Length 835;
Best Local Similarity 70.6%; Pred. No. 7.3e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 3 LPPQWEERQDILGRTYYVNHESRRRTQMKRTPPD 36
| | | | | | | | | | | | | | | | | | | |
Db 74 LPPQWEKVDNLGRTYYVNHNRSTQWHRPSLMD 107
| | | | | | | | | | | | | | | | | | | |

RESULT 8
Q9H2W4 PRELIMINARY; PRT; 854 AA.
ID Q9H2W4 AC Q9H2W4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NEDD4L (Ubiquitin ligase NEDD4Lb) (Ubiquitin ligase NEDD4f).
GN Name=NEDD4L; Synonyms=NEDD4L;
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21828840; PubMed=11840194; DOI=10.1038/sj.ejhg.5200747;
RA Chen H., Ross C.A., Wang N., Huo Y., MacKinnon D.F., Potash J.B.,
RA Simpson S.G., McMahon F.J., DePaulo J.R., Jr., McInnis M.G.;
RT "NEDD4L on human chromosome 18q21 has multiple forms of transcripts
RT and is a homologue of the mouse Nedd4-2 gene.";
RL Eur. J. Hum. Genet. 9:922-930(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009;
RA Qi H., Grenier J., Fournier A., Labrie C.;
RT "Androgens differentially regulate the expression of NEDD4L
RT transcripts in LNCaP human prostate cancer cells.";
RL Mol. Cell. Endocrinol. 210:51-62(2003).
DR EMBL; AF210730; AAG43524.1; -
DR EMBL; AF395931; AAM46208.1; -
DR EMBL; AY112983; AAM76728.1; -
DR HSP; Q62940; 115H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PRO0403; WNDOMAIN.
DR SMART; SM00119; HECTC; 1.

```

```
DR SMART; SM00456; WW; 4.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 4.
KW Ligase.
SQ SEQUENCE 854 AA; 98180 MW; 00C74E1661F52E7F CRC64;

Query Match 65.0%; Score 141; DB 2; Length 854;
Best Local Similarity 70.6%; Pred. No. 7.5e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDQ 36
|||||:|||||:|||||:|||||:
Db 74 LPPGWEKVDNLGRYYVNHNNRRTQWHRPSLMD 107

RESULT 9
Q9BW58 PRELIMINARY; PRT; 855 AA.
AC Q9BW58; Q8WUJ9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NEDD4L protein (Hypothetical protein) (Fragment).
GN Name=NEDD4L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Haie H.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshyuki S., Carninci F., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX Strausberg R.;
RA Strausberg R.;
RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX Strausberg R.;
RA Strausberg R.;
RN Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC000621; AAH00621.2; -.
DR EMBL; BC019345; AAH19345.1; -.
DR HSSP; Q62940; 115H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008973; C2 CALB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.

Query Match 65.0%; Score 141; DB 2; Length 855;
Best Local Similarity 70.6%; Pred. No. 7.5e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDQ 36
|||||:|||||:|||||:|||||:
Db 95 LPPGWEKVDNLGRYYVNHNNRRTQWHRPSLMD 128

RESULT 10
Q99PK2 PRELIMINARY; PRT; 855 AA.
AC Q99PK2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-protein ligase Nedd4-2.
GN Name=Nedd4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 Ncr1 BR;
RX MEDLINE=21067027; PubMed=11149908;
RA Kamyhina E., Debonneville C., Bens M., Vandewalle A., Staub O.;
RT "A novel mouse Nedd4 protein suppresses the activity of the epithelial
RT Na+ channel."
RL FASEB J. 15:204-214 (2001).
RL EMBL; AF277232; AAK00809.1; -.
DR HSSP; Q62940; 115H.
DR MGD; MGI:1913754; Nedd4l.
DR GO; GO:0005822; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 4.
KW Ligase.
SQ SEQUENCE 855 AA; 98258 MW; 3BE1566E2C3BA20E CRC64;

Query Match 65.0%; Score 141; DB 2; Length 855;
Best Local Similarity 70.6%; Pred. No. 7.5e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEERQDILGRYYVNHESRRRTQWKRPDQ 36
|||||:|||||:|||||:|||||:
Db 74 LPPGWEKVDNLGRYYVNHNNRRTQWHRPSLMD 107

RESULT 11
```

	Q8CFI0	PRELIMINARY;	PRT:	855 AA.				
ID	Q8CFI0							
AC	Q8CFI0							
DT	01-MAR-2003	(TrEMBLrel. 23, Created)						
DT	01-WAR-2003	(TrEMBLrel. 23, Last sequence update)						
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)						
DE	Neural cell expressed, developmentally down-regulated gene 4-like.							
GN	Name=Nedd4l;							
OS	Mus musculus (Mouse).							
OC	Eukaryota; Metazoa;							
OC	Mammalia; Eutheria; Rodentia;							
OX	NCBTaxId=10090;							
[1]								
RN	SEQUENCE FROM N.A.							
RP	STRAIN=FVB/N;	TISSUE=Kidney, and Mammary tumor. C3;						
RC	MEDLINE=22388257;	PubMed=12477932; DOI=10.1073/pnas.242603899;						
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,							
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Yoshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,							
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,							
RA	Fahey J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grichamwood J., Schmitz J., Myers R.M., Butterfield Y.S.,							
RA	Krzywinski M.I., Skalska U., Smalish D.E., Schnertch A., Schein J.E.,							
RA	Jones S.J., Marra M.A.;							
RT	Generation and initial analysis of more than 15,000 full-length human							
RL	and mouse cDNA sequences."							
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).							
[2]								
RN	SEQUENCE FROM N.A.							
RP	STRAIN=FVB/N;	TISSUE=Mammary tumor. C3;						
RC	Strausberg R.;							
RA	Submitted (NOV-2002) to the ENBL/GenBank/DBJ databases.							
[3]								
RN	SEQUENCE FROM N.A.							
RP	STRAIN=FVB/N;	TISSUE=Kidney;						
RC	Strausberg R.;							
RA	Submitted (JUN-2004) to the ENBL/GenBank/DBJ databaes.							
RL	EMBL; BC039746; AAAH39746.1; -							
DR	EMBL; BC071210; AAAH71210.1; -							
DR	HSP; O62940; IISH.							
DR	MGI; 1913754; Nedd4l.							
DR	GO; GO:0005622; C,intracellular; IEA.							
DR	GO; GO:0004842; F,ubiquitin-protein ligase activity; IEA.							
DR	GO; GO:0006512; P,ubiquitin cycle; IEA.							
DR	InterPro; IPR000569; HEAT.							
DR	InterPro; IPR002349; WW.							
DR	InterPro; IPR001202; WW_Rsp5_WWP.							
DR	Pfam; PF00632; HECT; 1.							
DR	Pfam; PF00397; WW; 4.							
DR	PRINTS; PR00403; WWDOMAIN.							
DR	SMART; SM00119; HECTG; 1.							
DR	SMART; SM00456; WW; 4.							
DR	PROSITE; PS00237; HECT; 1.							
DR	PROSITE; PS01159; WW_DOMAIN_1; 4.							
DR	PROSITE; PSS0020; WW_DOMAIN2; 4.							
SQ	SEQUENCE 855 AA; 98465 MW; 96C452B44285895 CRC64;							
Query Match                65.0%; Score 141; DB 2; Length 855;								
Best Local Similarity     70.6%; Pred.No. 7.5e+10;								
Matches      24; Conservative      3; Mismatches      7; Indels          0; Gaps								
QY	3 LPFGWEERDILGRYYYNHESRRTPQWRKPTPD 36      ...							

```

MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny N.B., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 C2 domain.
DR EMBL; BC032597; ANH32597.1; -.
DR HSSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
SQ SEQUENCE 911 AA; 104921 MW; CE04AAED677AA506 CRC64;

Query Match 65.0%; Score 141; DB 2; Length 911;
Best Local Similarity 70.6%; Pred. No. 8e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEEQDILGRYYVNHESRRRTQWKRTPOD 36
Db 195 LPPGWEEKVDNLGRYYVNHNNRTTQWHRPSLMD 228

RESULT 14
Q7Z5F1
ID Q7Z5F1 PRELIMINARY; PRT; 947 AA.
AC Q7Z5F1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin ligase NEDD4h.
GN Name=NEDD4L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Qy 3 LPPGWEEQDILGRYYVNHESRRRTQWKRTPOD 36
Db 195 LPPGWEEKVDNLGRYYVNHNNRTTQWHRPSLMD 228

RESULT 14
Q7Z5F1
ID Q7Z5F1 PRELIMINARY; PRT; 947 AA.
AC Q7Z5F1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin ligase NEDD4h.
GN Name=NEDD4L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny N.B., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 C2 domain.
DR EMBL; BC032597; ANH32597.1; -.
DR HSSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
SQ SEQUENCE 911 AA; 104921 MW; CE04AAED677AA506 CRC64;

Query Match 65.0%; Score 141; DB 2; Length 911;
Best Local Similarity 70.6%; Pred. No. 8e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEEQDILGRYYVNHESRRRTQWKRTPOD 36
Db 195 LPPGWEEKVDNLGRYYVNHNNRTTQWHRPSLMD 228

RESULT 14
Q7Z5F1
ID Q7Z5F1 PRELIMINARY; PRT; 947 AA.
AC Q7Z5F1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin ligase NEDD4h.
GN Name=NEDD4L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny N.B., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 C2 domain.
DR EMBL; BC032597; ANH32597.1; -.
DR HSSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
SQ SEQUENCE 911 AA; 104921 MW; CE04AAED677AA506 CRC64;

Query Match 65.0%; Score 141; DB 2; Length 947;
Best Local Similarity 70.6%; Pred. No. 8.4e-10;
Matches 24; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LPPGWEEQDILGRYYVNHESRRRTQWKRTPOD 36
Db 187 LPPGWEEKVDNLGRYYVNHNNRTTQWHRPSLMD 220

RESULT 15
Q96PU5
ID Q96PU5 PRELIMINARY; PRT; 955 AA.
AC Q96PU5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NEDD4-like ubiquitin ligase 3.
GN Name=NEDL3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20501262; PubMed=11046148;
RX DOI=10.1126/MCB.20.22.8526-8535.2000;
RA Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G.,
RA Ingham R., Ernsberg I., Pawson T.;
RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";
RL Mol. Cell. Biol. 20:8526-8535 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269431; PubMed=11244092; DOI=10.1074/jbc.C000906200;
RX Harvey F.K., Dinudom A., Cook I.D., Kumar S.;
RT "The Nedd4-like protein KIAA0439 is a potential regulator of the
RT epithelial sodium channel.";
RL J. Biol. Chem. 276:8597-8601 (2001).
RN [3]
RP SEQUENCE FROM N.A.
```





**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 85.6044 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-5  
Perfect score: 221  
Sequence: 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_15Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	221	100.0	38	4	AAB83020 Human Ned
2	221	100.0	38	7	ADB49217 Human WW
3	221	100.0	733	6	AAB32726 Human NED
4	221	100.0	900	8	ADQ07991 Human NED
5	221	100.0	927	2	AAY25170 Human KIA
6	221	100.0	927	6	AAB32725 Human NED
7	187	84.6	38	5	AAU87973 Human WW
8	187	84.6	38	7	ADB49327 Novel hum
9	187	84.6	724	2	AAB36797 Novel hum
10	187	84.6	725	7	ADB49320 Novel hum
11	187	84.6	818	2	AAW13386 Human pro
12	187	84.6	830	8	ADQ97134 Human can
13	187	84.6	834	5	AAU77715 Diseased
14	187	84.6	854	6	ABP58332 Human cel
15	187	84.6	854	7	ADB75453 Prostate
16	187	84.6	911	7	ADB75455 Prostate
17	187	84.6	923	4	ABG01080 Novel hum
18	187	84.6	940	8	ADB66480 Human pro
19	187	84.6	940	8	ADB66138 Human pro
20	187	84.6	941	8	ADQ97130 Human can
21	187	84.6	948	8	ADQ55161 Protein #
22	187	84.6	954	8	ADQ97125 Mouse can
23	187	84.6	955	8	ADQ07990 Human NED
24	187	84.6	958	8	ADQ97128 Human can
25	187	84.6	975	2	AAW93167 Human ZGG

ALIGNMENTS

RESULT 1

AAAB83020					
ID	AAAB83020	standard; peptide; 38 AA.			
XX	AC	AAAB83020;			
XX	DT	25-JUN-2001 (first entry)			
XX	DE	Human Nedda HECT E3 ubiquitin ligase WW domain #3.			
XX	KW	Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;			
KW	HECT E3 ubiquitin ligase; antiinflammatory; cytosolic; osteopachic;				
KW	antimicrobial; neuroprotective; transforming growth factor beta;				
KW	TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;				
KW	inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;				
KW	renal failure; neurodegeneration; fibrosis; WW domain; Nedda.				
XX	OS	Homo sapiens.			
XX	PN	WO200116604-A1.			
XX	PD	08-MAR-2001.			
XX	PF	29-AUG-2000; 2000WO-US023729.			
XX	PR	30-AUG-1999; 99US-00385918.			
XX	PA	(SIGN-) SIGNAL PHARM INC.			
XX	PI	Hoekstra MF, Xie W, Murray BW, Mercurio FM;			
XX	DR	WPI; 2001-327913/34.			
PT	PT	Screening for modulators of TGF-beta and/or bone morphogenic protein			
PT	PT	(BMP) mediated signaling useful for treating cancer and osteoporosis by			
XX	XX	evaluating the ability of agents to modulate Smad protein degradation.			
XX	XX	Disclosure; Page 13; 75pp; English.			
CC	CC	The present sequence is the WW domain of a HECT (homologous to E6			
CC	CC	carboxyl terminus) E3 ubiquitin ligase. The WW domain binds to the Smad			
CC	CC	py motif, resulting in ubiquitination of Smad by the E3 ubiquitin ligase.			
CC	CC	The sequence is provided in a specification relating to a method for			
CC	CC	screening for agents that modulate transforming growth factor (TGF)-beta			
CC	CC	and/or bone morphogenic protein (BMP)-mediated signalling. The method			
CC	CC	involves evaluating the effect of an agent on binding of HECT E3			
CC	CC	ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad			
CC	CC	protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein			

CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
 CC bone formation in a patient or treating a condition associated with  
 CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
 CC inhibit BMP-mediated signalling are useful for treating inflammation,  
 CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
 CC signalling are useful for stimulating bone anabolism as well as treating  
 CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
 CC that inhibit TGF-mediated signalling are useful for treating cancer,  
 CC inflammation, neurodegeneration and fibrosis

XX Sequence 38 AA;

Query Match 100.0%; Score 221; DB 4; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-22;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||  
 Db 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||

RESULT 2  
 ADB49217

ID ADB49217 standard; peptide; 38 AA.

XX AC ADB49217;

DT 04-DEC-2003 (first entry)

XX Human WW binding protein Ned4 WW domain #3.

DE WW domain; drug candidate screening; drug discovery; drug modification;  
 XX drug refinement; immunogen; WW binding protein; human; ned4; WW domain.  
 KW Homo sapiens.

OS US2003077577-A1.

PN 24-APR-2003.

XX 28-JUN-2002; 2002US-00185050.

XX 03-APR-1996; 96US-00630916.

PR 03-APR-1997; 97US-00826516.

XX (PIRO/) PIROZZI G.

PA (KAYB/) KAY B K.

PA (FOWL/) FOWLKES D M.

XX Pirozzi G, Kay BK, Fowlkes DM;  
 WPI; 2003-635075/60.

XX Novel purified polypeptide comprising WW domain, useful for drug  
 PT discovery, modification and refinement, for discovering polypeptides  
 PT involved in pharmacological activities, or as an immunogen to generate  
 PT antibodies.

XX Example; Fig 5; 133pp; English.

CC The invention describes a purified polypeptide (I) comprising a WW domain  
 CC which has a sequence (S1) selected from 11 sequences fully defined in the  
 CC specification, a sequence (S2) selected from 48 sequences fully defined  
 CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725  
 CC amino acids fully defined in the specification. (I) is useful for  
 CC screening a potential drug candidate, by allowing (I) to come into  
 CC contact with at least one recognition unit having a selective affinity  
 CC for the WW domain in (I), in the presence of an amount of a potential  
 CC drug candidate, such that (I) and the recognition unit are capable of  
 CC interacting when brought into contact with one another in the absence of  
 CC the drug candidate, and determining the effect, if any, of the presence  
 CC of the amount of the drug candidate on the interaction of (I) with the  
 CC recognition unit. (I) is useful for drug discovery, modification and

CC refinement, for discovering polypeptides involved in pharmacological  
 CC activities, or as an immunogen to generate antibodies. This is the amino  
 CC acid sequence of WW binding protein Ned4 WW domain.

XX Sequence 38 AA;

Query Match 100.0%; Score 221; DB 7; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-22;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||  
 Db 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||

RESULT 3  
 AAE32726

ID AAE32726 standard; protein; 733 AA.

XX AC AAE32726;

XX 24-MAR-2003 (first entry)

XX Human NEDD4 short form protein.

DE Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;  
 XX poliomyelitis; HIV; measles; protein therapy; human; NEDD4.

KW Homo sapiens.

OS

XX Key Location/Qualifiers  
 FH Misc-difference 260 /note= "Encoded by CAA"  
 FT Misc-difference 279 /note= "Encoded by AGC"  
 FT WO200209549-A2.

XX 14-NOV-2002.

XX 12-MAR-2002; 2002WO-IB002106.

XX 12-MAR-2001; 2001US-0275224P.

PR 31-JUL-2001; 2001US-0308958P.

PR 07-DEC-2001; 2001US-0340170P.

XX (PROT-) PROTEOLOGICS LTD.

XX Greener T, Moskowitz H, Reiss Y, Alroy I;  
 WPI; 2003-111976/10.  
 DR N-PSDB; AAD50460.

XX New protein complex comprising HECT-RC1, viral maturation scaffolding  
 PT protein (VMSp), and/or HIV gag protein, useful for treating viral  
 PT infections, such lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,  
 PT or Ebola.

XX Disclosure; Fig 16; 150pp; English.

CC The invention relates to a method for modulation of viral maturation. The  
 CC invention also provides an isolated protein complex comprising a HECT-  
 CC RC1 selected from HECT-WW, HECT-RC1, Gag protein, Gag late domain, P13,  
 CC actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIIM,  
 CC GTPase, E2 enzyme, tsg101, cullin, HECT1, HECT2, HECT3, Nedd4-like  
 CC protein or clathrin. The complexes, proteins, antibodies and methods are  
 CC useful for treating viral infections, such as lymphosarcoma, human  
 CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola  
 CC and for inhibiting budding in a subject. They are also useful in  
 CC diagnostic assays for determining whether a cell is infected with a virus  
 CC and for characterizing the nature, progression and/or infectivity of the  
 CC infection. The invention is also useful in protein therapy. The present  
 CC sequence is human NEDB4 protein used to illustrate the method of the

CC invention  
XX Sequence 733 AA;  
SQ

Query Match 100.0%; Score 221; DB 6; Length 733;  
Best Local Similarity 100.0%; Pred. No. 6.7e-21;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWDPRLKIPA 38  
|||||  
Db 421 GFLPKGWEVRHAPNGRPFIDHNTKTTTWDPRLKIPA 458  
|||||

RESULT 4  
ADQ07991  
ID ADQ07991 standard; protein; 900 AA.  
XX  
AC ADQ07991;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human NEDD4 protein.  
XX  
KW MBGAT: modifier of beta-catenin; drug screening;  
KW beta-catenin pathway modulator; wingless signalling pathway;  
KW wg signalling pathway; wnt signalling pathway; cell cycle progression;  
KW diagnosis; cancer; angiogenic disorder; apoptotic disorder;  
KW cell proliferation disorder; cytostatic; gene therapy; human; NEDD4;  
KW neural precursor cell expressed developmentally downregulated 4.  
XX  
OS Homo sapiens.  
XX  
XX WO2004061123-A2.  
XX  
XX 22-JUL-2004.  
XX  
XX 29-DEC-2003; 2003WO-US041404.  
XX  
XX 30-DEC-2002; 2002US-0436974P.  
XX  
XX (EXEL-) EXELIXIS INC.  
XX  
XX Francis-Lang H, Winter CG, Ventura RBA, Lickteig K;  
XX  
XX WPI; 2004-534389/51.  
XX  
XX N-PSDB; ADQ07980.  
XX  
XX REFSEQ; XM\_046129.4.  
XX  
XX Identifying a candidate beta-catenin MBGAT pathway modulating agent,  
XX useful for diagnosing and treating cancer, comprises providing an assay  
XX system comprising a MBGAT polypeptide or nucleic acid.  
XX  
XX Example II; SEQ ID NO 22; 113pp; English.  
XX  
XX The invention relates to the use of MBGAT (modifier of beta-catenin)  
XX polypeptides or nucleic acids for identifying a candidate beta-catenin  
XX pathway modulating agent or for modulating a beta-catenin pathway of a  
XX cell. The invention also relates to a method of modulating a beta-catenin  
XX pathway in a cell; and diagnosing a disease in a patient by detecting  
XX MBGAT expression. The invention is based on the finding that MBGAT  
XX orthologues are able to modify the beta-catenin pathway in Drosophila  
XX cells. Beta-catenin (also known as Armadillo in Drosophila) plays a role  
XX in cell signalling via the Wntless (Wg)/Wnt signalling pathway which is  
XX involved in cell cycle progression. Beta-catenin levels are tightly  
XX regulated by a complex containing the tumour suppressor protein APC  
XX (adenomatous polyposis coli) and Axin and the serine/threonine kinase  
XX GSK3-beta. In the absence of these proteins, beta-catenin accumulates in  
XX the nucleus where it acts as a transcriptional co-activator with TCF for  
XX the induction of target genes such as the cell-cycle regulators cyclin D1  
XX and c-Myc. The Wntless/beta-catenin signalling pathway is frequently  
XX mutated in human cancers, particularly those of the colon. Beta-catenin  
XX pathway modulating agents identified using the methods of the invention  
XX are useful as therapeutic agents for treating disorders associated with

CC defective or impaired beta-catenin function and/or MBGAT function such as  
CC cancer, or angiogenic, apoptotic or cell proliferation disorders.  
CC proteins that interact with MBGAT are also useful in detecting and  
CC providing information about the function of MBGAT proteins. The present  
CC sequence represents a human NEDD4 (neural precursor cell expressed,  
CC developmentally downregulated 4) protein, identified as an MBGAT  
CC polypeptide in an example of the invention.  
XX  
SQ Sequence 900 AA;  
Query Match 100.0%; Score 221; DB 8; Length 900;  
Best Local Similarity 100.0%; Pred. No. 8.5e-21;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWDPRLKIPA 38  
|||||  
Db 421 GFLPKGWEVRHAPNGRPFIDHNTKTTTWDPRLKIPA 458  
|||||

RESULT 5  
AAAY25170  
ID AAAY25170 standard; protein; 927 AA.  
XX  
AC AAAY25170;  
XX  
DT 09-SEP-1999 (first entry)  
XX  
DE Human KIAAN ligase protein fragment.  
XX  
KW Inhibitor; ubiquitin-mediated proteolysis; IkappaB polypeptide; HECT;  
KW ubiquitin-conjugating system; Homologous to E6-Ap carboxyl terminus;  
KW ligase; ubiquitin; ubiquitination; RSC ligase; KIAAN ligase; UBC4; MAD3;  
KW modulator; treatment; proliferative disorder; apoptosis; sepsis;  
KW differentiative disorder; viral infection; tissue wasting disorder;  
KW cachexia; malignancy; inflammatory disease; parasitic disease;  
KW tuberculosis; IL-2 therapy; rheumatoid arthritis; rheumatoid spondylitis;  
KW osteoarthritis; gouty arthritis; respiratory distress syndrome;  
KW cerebral malaria; chronic pulmonary inflammatory disease; silicosis;  
KW pulmonary sarcoidosis; bone resorption disease; reperfusion injury;  
KW graft versus host reaction; allograft rejection; Crohn's disease;  
KW ulcerative colitis; pyresis; multiple sclerosis; autoimmune diabetes;  
KW systemic lupus erythematosus; leprosy; AIDS.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Protein 1..927  
XX /note= "partial sequence, no start codon given"  
XX  
XX WO9904033-A1.  
XX  
XX 28-JAN-1999.  
XX  
XX 16-JUL-1998; 98WO-US014638.  
XX  
XX 16-JUL-1997; 97US-00895601.  
XX  
XX (MITO-) MITOTIX INC.  
XX  
XX Beer-Romero P, Strack PR, Glass SJ, Rolfe M;  
XX WPI; 1999-132274/11.  
XX  
XX N-PSDB; AAX78494.  
XX  
XX Identifying modulators of I-kappa-B proteolysis - used to develop  
XX products for treating e.g. proliferative and/or differentiative  
XX disorders, infections, tissue wasting, cachexia or AIDS.  
XX  
XX Disclosure; Page 68-71; 79pp; English.  
XX  
XX This invention describes a novel assay identifying an inhibitor of  
XX ubiquitin-mediated proteolysis of a IkappaB polypeptide. The method  
XX comprises: (a) providing a ubiquitin-conjugating system including the

CC IkappaB polypeptide and a HECT (Homologous to E6-Ap Carboxyl Terminus)  
 CC ligase and ubiquitin, to promote ubiquitination of the IkappaB  
 CC polypeptide by the HECT ligase; (b) contacting the ubiquitin-conjugating  
 CC system with a candidate agent; (c) measuring a level of ubiquitination of  
 CC the polypeptide in the presence of the candidate agent and (d) comparing  
 CC the measured level of ubiquitination in the presence of the candidate  
 CC agent with ubiquitination of the IkappaB polypeptide in the absence of  
 CC the candidate agent; where a statistically significant decrease in  
 CC ubiquitination of the IkappaB polypeptide in the presence of the  
 CC candidate agent is indicative of an inhibitor of ubiquitination of the  
 CC IkappaB polypeptide. The assays can be used to identify compounds which  
 CC modulate binding and/or ubiquitination of an IkappaB (or other cellular  
 CC or viral substrate) by a HECT ligase, such as RSC or K1AAN. Such  
 CC modulators can be used e.g. in the treatment of proliferative and/or  
 CC differentiative disorders, to modulate apoptosis, in the treatment of  
 CC viral infections, and in the treatment of tissue wasting disorders e.g.  
 CC cachexia secondary to infection or malignancy, cachexia secondary to  
 CC human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and  
 CC high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis,  
 CC osteoarthritis, gouty arthritis and other arthritic conditions, sepsis,  
 CC respiratory distress syndrome, cerebral malaria, chronic pulmonary  
 CC inflammatory disease, silicosis, pulmonary sarcoidosis, bone resorption  
 CC diseases, reperfusion injury, graft versus host reaction, allograft  
 CC rejections, Crohn's disease, ulcerative colitis, or pyresis, in addition  
 CC to a number of autoimmune diseases such as multiple sclerosis, autoimmune  
 CC diabetes, systemic lupus erythematosus; and ENL in leprosy, HIV, and  
 CC AIDS. This sequence represents a human K1AAN ligase which is used in the  
 CC method of the invention  
 XX  
 SQ Sequence 927 AA;

Query Match 100.0%; Score 221; DB 2; Length 927;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-21;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||  
 Db 448 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 485

RESULT 6  
 AAEE32725  
 ID AAE32725 standard; protein; 927 AA.

XX  
 AC AAE32725;

XX 24-MAR-2003 (first entry)

XX Human NEDD4 long form protein.

XX Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;  
 KW poliomylitis; HIV; measles; protein therapy; human; NEDD4.

XX Homo sapiens.

XX Key Location/Qualifiers  
 PH 381..406  
 FT Domain /note= "WW domain"  
 FT Domain 821..923  
 FT /note= "HECT domain"

XX WO200290549-A2.

XX 14-NOV-2002.

XX 12-MAR-2002; 2002WO-IB002106.

XX 12-MAR-2001; 2001US-0275224P.

XX 31-JUL-2001; 2001US-0308958P.

XX 07-DEC-2001; 2001US-0340170P.

XX (PROT-) PROTEOLOGICS LTD.

XX

PI Greener T, Moskowitz H, Reiss Y, Alroy I;  
 XX WPI; 2003-111976/10.  
 DR N-PSDB; AAD50459.  
 XX  
 PT New protein complex comprising HECT-RC1, viral maturation scaffolding  
 PT protein (VMSp), and/or HIV gag protein, useful for treating viral  
 PT infections, such lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,  
 PT or Ebola.  
 XX  
 XX Disclosure; Fig 15; 150pp; English.  
 PS  
 XX The invention relates to a method for modulation of viral maturation. The  
 XX invention also provides an isolated protein complex comprising a HECT-  
 CC RC1 selected from HECT-WW, HECT-RC1, Gag protein, Gag late domain, P13,  
 CC actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM,  
 CC GPase, E2 enzyme, tsq101, cullin, HECT1, HECT2, HECT3, Nedd4-like  
 CC protein or clathrin. The complexes, proteins, antibodies and methods are  
 CC useful for treating viral infections, such as lymphosarcoma, human  
 CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola  
 CC and for inhibiting budding in a subject. They are also useful in  
 CC diagnostic assays for determining whether a cell is infected with a virus  
 CC and for characterizing the nature, progression and/or infectivity of the  
 CC infection. The invention is also useful in protein therapy. The present  
 CC sequence is human NEDD4 protein used to illustrate the method of the  
 CC invention  
 XX  
 SQ Sequence 927 AA;

Query Match 100.0%; Score 221; DB 6; Length 927;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-21;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38  
 |||||  
 Db 448 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 485

RESULT 7  
 AAU87973

ID AAU87973 standard; peptide; 38 AA.

XX AAU87973;

XX 05-JUN-2002 (first entry)

XX Human WW domain #11.

XX Human; PDZ domain; WW domain; rat; cow; mouse; fruitfly; protein therapy;  
 KW gene therapy; PDZ-mediated disease; inward potassium channel; WBP;  
 KW dimer inhibitor peptide; carboxylate binding loop.

XX Homo sapiens.

XX WO200207751-A1.

XX 31-JAN-2002.

XX 24-JUL-2001; 2001WO-US023269.

XX 25-JUL-2000; 2000US-0221215P.

XX 28-NOV-2000; 2000US-00723810.

XX (AXCE-) AXCELL BIOSCIENCES CORP.

XX Herrero J, Pirozzi G, Uveges A;

XX WPI; 2002-195842/25.

XX Methods for identifying polypeptides comprising PDZ domains, the  
 PT polypeptides and their encoding nucleic acids, useful for the diagnosis  
 PT and treatment of PDZ related disorders.

XX

PS Disclosure; Fig 20; 225pp; English.

XX The invention relates to methods for identifying polypeptides comprising

CC PDZ domains, and their encoding nucleic acids. The sequences are used to

CC identify modulators of their expression, function and activity, for use

CC in the diagnosis and treatment of PDZ related disorders. Antibodies

CC against the proteins and cells that produce them may be used for the

CC treatment of PDZ-mediated disease states. Sequences AAU8743-AAU87974

CC represent proteins containing PDZ domains, fragments of these proteins

CC and other related peptides used in the methods of the invention

XX SQ Sequence 38 AA;

Query Match 84.6%; Score 187; DB 5; Length 38;

Best Local Similarity 88.9%; Pred. No. 8.8e-18;

Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FLPGWVVRHAPNGRPFIDHNTKTTTWEDEPRLKIP 37

Db 2 FLPPGWEMRIAPNGRPFIDHNTKTTTWEDEPRLKFP 37

RESULT 8

ADB49327

ID ADB49327 standard; peptide; 38 AA.

XX AC

AC ADB49327;

DT 04-DEC-2003 (first entry)

XX DE Novel human WWP4-2 protein WW domain.

XX KW WW domain; drug candidate screening; drug discovery; drug modification;

XX KW drug refinement; immunogen; WW binding protein 4; WWP4; human.

XX OS Homo sapiens.

XX US2003077577-A1.

XX PD 24-APR-2003.

XX PF 28-JUN-2002; 2002US-00195050.

XX PR 03-APR-1996; 96US-00630916.

XX PR 03-APR-1997; 97US-00826516.

XX PA (PIROZZI) PIROZZI G.

XX PA (KAYB/) KAY B K.

XX PA (FOWL/) FOWLKES D M.

XX PI Pirozzi G, Kay BK, Fowlkes DM;

XX WPI; 2003-635075/60.

XX Novel purified polypeptide comprising WW domain, useful for drug

PT discovery, modification and refinement, for discovering polypeptides

PT involved in pharmacological activities, or as an immunogen to generate

PT antibodies.

XX Example; Fig 5; 133pp; English.

XX The invention describes a purified polypeptide (I) comprising a WW domain

CC which has a sequence (S1) selected from 11 sequences fully defined in the

CC specification, a sequence (S2) selected from 48 sequences fully defined

CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725

CC amino acids fully defined in the specification. (I) is useful for

CC screening a potential drug candidate, by allowing (I) to come into

CC contact with at least one recognition unit having a selective affinity

CC for the WW domain in (I), in the presence of an amount of a potential

CC drug candidate, such that (I) and the recognition unit are capable of

CC interacting when brought into contact with one another in the absence of

CC the drug candidate, and determining the effect, if any, of the presence

CC of the amount of the drug candidate on the interaction of (I) with the

CC

CC recognition unit. (I) is useful for drug discovery, modification and

CC refinement, for discovering polypeptides involved in pharmacological

CC activities, or as an immunogen to generate antibodies. This is the amino

CC acid sequence of novel human WW binding protein WWP4-2 WW domain.

XX SQ Sequence 38 AA;

Query Match 84.6%; Score 187; DB 7; Length 38;

Best Local Similarity 88.9%; Pred. No. 8.8e-18;

Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FLPGWVVRHAPNGRPFIDHNTKTTTWEDEPRLKIP 37

Db 2 FLPPGWEMRIAPNGRPFIDHNTKTTTWEDEPRLKFP 37

RESULT 9

AAW36797

ID AAW36797 standard; peptide; 724 AA.

XX AC

AC AAW36797;

DT 23-APR-1998 (first entry)

XX DE Novel human gene, designated WWP4.

XX KW Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;

XX KW WW domain; cell signalling; growth regulation; cytoskeleton organisation;

XX KW targeted drug screening; modulator; WW domain interaction; WWP4.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT misc\_difference 1.3

FT /note= "the nucleotides encoding these amino acids are

FT not given in the specification"

FT Domain 140..165

FT /note= "claimed (claim 49) WW Domain 1"

FT Domain 252..277

FT /note= "claimed (claim 49) WW domain 3"

FT Domain 303..328

FT /note= "claimed (claim 49) WW domain 3"

FT Domain 618..724

FT /note= "claimed (claim 90) HECT domain"

XX WO9737223-A1.

XX 09-OCT-1997.

XX 03-APR-1997; 97WO-US005547.

XX 03-APR-1996; 96US-00630916.

XX (CYTO-) CYTOGEN CORP.

XX (UNNC-) UNIV NORTH CAROLINA.

XX Pirozzi G, Kay BK, Fowlkes DM;

XX WPI; 1997-503234/46.

XX N-PSDB; AAT95700.

XX Identifying cell signalling and growth regulatory polypeptides by

PT reaction with multivalent recognition complex - polypeptides are useful

PT in targetted drug selection.

XX Claim 48; Fig 23; 220pp; English.

XX The present sequence represents a novel protein WWP4. The WWP4 gene was

CC identified and isolated from a cDNA expression library generated from

CC LNCap prostate cancer cell line, using peptides AAW38063-64. These

CC peptide recognition units are based on the sequences of WW domain binding

CC domains of the alpha and gamma subunits of epithelial sodium channel

CC protein. The WW domain is a small functional domain found in a large

CC





Best Local Similarity 88.9%; Pred. NO. 3e-16;  
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

**Qy**                2   FLPKGWEVRHAPNGRPFIDHNKTTTWTWEDPRUKIP   37  
                 ||| ||| : | ||||| ||||| ||||| |||||  
**D6**                357   FLPPGWEVRHAPNGRPFIDHNKTTTWTWEDPRUKEP   392

RESULT 12  
ADQ97134  
ID ADQ97134 standard: protein: 830 AA.

XX  
AC  
ADQ97134;  
XX  
XX  
XX  
DT  
DT  
XX  
XX  
DE  
DE  
XX  
XX  
KW  
KW  
Human cancer associated sequence HP4-08-005, SEQ ID 110.  
Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.

AA  
OS  
Homo sapiens.

XX PN WO2004060304-A2.

XX  
PD 22-JUL-2004

XX 22-DEC-2003: 2003WO-US041389.

XX  
PR 27-DEC-2002: 2002US-00330773.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX  
PI  
PI  
Morris DW, Malandro MS;

XX  
DR  
WPI: 2004-543781/52.

PT New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.

PS Claim 1: SEO ID NO 110: 199pp: English.

The present invention relates to cancer associated sequences (ADQ97025-ADQ9804). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://www.int/pub/published/nct/sequences>

XX  
SQ  
Sequence 830 AA:

Query Match 84.6%; Score 187; DB 8; Length 830;  
Best Local Similarity 88.9%; Pred. No. 3.1e-16;  
Matches 12: Conservative 1; Mismatches 3; Indels

Qy 2 FLPKGWEVRHANGRPFFIDHNKTWTTWEDPRUKIP 37  
||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| |||  
n6 353 FTDDCGNMPIALNGCPDPTIDHNKATTTWTFNDOKCD 189

RESULT 13  
AAU77715  
ID AAU77715 standard: protein: 834 AA

XX  
AC 281177715.

XX  
DT 05-JUN-2002 (first entry)

DE Diseased kidney tissue associated novel human protein #2

XX Kidney disease; nephrotropic; renal disease; IgA nephropathy; IgAN;  
KW asymptomatic microscopic haematuria; proteinuria;  
KW episodic gross haematuria; necrotising crescentic glomerulonephritis;  
XX

NCGN; minimal change disease; Min Ch; renal hypertrophy; diabetes;  
 end-stage renal disease; ESRD; kidney function.

Homo sapiens.  
WO200212439-A2.

14-FEB-2002.

03-AUG-2001; 2001WO-US024635.

03-AUG-2000; 2000US-0222731P.

(GENE-) GENE LOGIC INC.

Sun H, Munger WE, Falk RJ;

WPI; 2002-257475/30.

N-PSDB; ABK11666:

NOVER human genes that are differentially expressed in renal biopsy samples from patients with kidney diseases useful for diagnosing renal diseases such as IgA nephropathy or necrotizing crescentic glomerulonephritis.

Claim 16: Page 62-65: 67pp: English:

The invention describes an isolated nucleic acid molecule (I) that is expressed in diseased kidney tissue compared to normal kidney tissue. (I) and its encoded protein (II) are useful for diagnosing a renal disease state e.g. IgA nephropathy (IgAN) which presents as asymptomatic microscopic haematuria and/or proteinuria or episodic gross haematuria, necrotising crescentic glomerulonephritis (NCGN), minimal change disease (Min Ch) or renal hypertrophy (the cause of end-stage renal disease and early manifestation of diabetes), in a subject. (II) is useful: as a target for identifying agents which modulate the expression or activity of (I); for identifying binding partners: as novel target for screening synthetic small molecules and combinatorial or naturally occurring compound libraries to discover novel therapeutics to regulate kidney function; as an antigen to raise polyclonal or monoclonal antibodies, and as a therapeutic agent or target. Agents that modulate, up- or down-regulate the expression of (II) or agents that agonise or antagonise the activity of the protein are useful to modulate biological and pathologic purposes associated with proteins function and activity, for e.g. kidney damage or end-stage renal disease (ESRD). This is the amino acid sequence of the diseased kidney associated novel human protein #2, described in the invention

Sequence 834 AA:

Query Match 84.6%; Score 187; DB 5; Length 834;  
Best Local Similarity 88.9%; Pred. No. 3.1e-16;  
Matches 32: Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy 2 FLPKGVRHAPNGRPFIDHNTKTTTWEDPRUKIP 37
      |||||:|||||
Db 157 FLPDGWRPIADNGRPFIDHNTKTTTWEDPRUKFP 192
```

RESULT 14  
ABP58332  
IN ABP58332 standard: protein: 854 AA

XX  
AC  
ABD58332.XX  
DT 07-APR-2003 (first entry)

Human cell growth differentiation and death protein CGDP-3

XX CGD-3; cell growth; cell differentiation; cell death; human; cytostatic;  
XX antiarteriosclerotic; hepatotropic; antiinflammatory; antipsoriatic;  
KW antianaemic; ophthalmological; auditory; anticonvulsant;  
KW cerebroprotective; neurotropic; neuroprotective; antiparkinsonian;  
KW





```
RESULT 2
US-08-895-601-6
; Sequence 6, Application US/08895601
; Patent No. 6060262
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Strack, Peter J.
; APPLICANT: Glass, Susan J.
; APPLICANT: Rolfe, Mark
; TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,
; AND METHODS AND REAGENTS RELATED THERETO
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895.601
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-096.01
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 927 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-895-601-6

Query Match 100.0%; Score 221; DB 3; Length 927;
Best Local Similarity 100.0%; Pred. No. 3,9e-22;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38
|||||
Db 448 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 485

RESULT 3
US-08-539-205A-6
; Sequence 6, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392.163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-163A-6
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/539.205A
; APPLICATION NUMBER: US/08/539.205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-6

Query Match 84.6%; Score 187; DB 3; Length 834;
Best Local Similarity 88.9%; Pred. No. 1.8e-17;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FLPGKWEVRHAPNGRPFIDHNTKTTTWEDPRLKIP 37
|||||
Db 357 FLPGKWEVRHAPNGRPFIDHNTKTTTWEDPRLKIP 392

RESULT 4
US-09-392-163A-6
; Sequence 6, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392.163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-163A-6
```

```

Query Match      84.6%; Score 187; DB 4; Length 834;
Best Local Similarity 88.9%; Pred. No. 1.8e-17;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FLPLKGEVHRHAPNGRPFIDHNTKTTTWEDPRLKIP 37
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
Db 357 FLPLPGWEMRIANGRPFIDHNTKTTTWEDPRLKIP 392

RESULT 5
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-539-205A-4

Query Match      66.1%; Score 146; DB 3; Length 766;
Best Local Similarity 69.7%; Pred. No. 8.1e-12;
Matches 23; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 GFLPKGEVHRHAPNGRPFIDHNTKTTTWEDPR 33
   ||| ||| : ||| ||| : ||| ||| : ||| |||
Db 288 GELPPGWEQRYTPGPRPVFDHNTTTTWDPR 320

RESULT 6
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA

```

```

; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392.163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-4

Query Match      66.1%; Score 146; DB 4; Length 766;
Best Local Similarity 69.7%; Pred. No. 8.1e-12;
Matches 23; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 GFLPKGEVHRHAPNGRPFIDHNTKTTTWEDPR 33
   ||| ||| : ||| ||| : ||| ||| : ||| |||
Db 288 GELPPGWEQRYTPGPRPVFDHNTTTTWDPR 320

RESULT 7
US-08-630-916A-22
; Sequence 22, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630.916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:

```

```
;
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-916A-22

Query Match 61.5%; Score 136; DB 3; Length 38;
Best Local Similarity 66.7%; Pred. No. 5.8e-12;
Matches 22; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNKTTTWEDPR 33
| || || | | | | | | | | | | | | | | | |
Db 1 GELPSGWEQRTPEGRAYFVDHNTRTTTWVDP 33
| || || | | | | | | | | | | | | | | | |

RESULT 8
US-08-348-518C-17
; Sequence 17, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-2
; IMMEDIATE SOURCE:
; CLONE: Rsp5
; US-08-348-518C-17

Query Match 61.5%; Score 136; DB 3; Length 38;
Best Local Similarity 66.7%; Pred. No. 5.8e-12;
Matches 22; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNKTTTWEDPR 33
| || || | | | | | | | | | | | | | | | |
Db 1 GELPSGWEQRTPEGRAYFVDHNTRTTTWVDP 33
| || || | | | | | | | | | | | | | | | |

RESULT 9
US-08-476-509B-17
; Sequence 17, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-2
; IMMEDIATE SOURCE:
; CLONE: Rsp5
; US-08-476-509B-17

Query Match 61.5%; Score 136; DB 3; Length 38;
Best Local Similarity 66.7%; Pred. No. 5.8e-12;
Matches 22; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNKTTTWEDPR 33
| || || | | | | | | | | | | | | | | | |
Db 1 GELPSGWEQRTPEGRAYFVDHNTRTTTWVDP 33
| || || | | | | | | | | | | | | | | | |

RESULT 10
US-09-270-767-59345
; Sequence 59345, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
Db 1 GELPSGWEQRTPEGRAYFVDHNTRTTTWVDP 33

RESULT 9
US-08-476-509B-17
; Sequence 17, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-2
; IMMEDIATE SOURCE:
; CLONE: Rsp5
; US-08-476-509B-17

Query Match 61.5%; Score 136; DB 3; Length 38;
Best Local Similarity 66.7%; Pred. No. 5.8e-12;
Matches 22; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFIDHNKTTTWEDPR 33
| || || | | | | | | | | | | | | | | | |
Db 1 GELPSGWEQRTPEGRAYFVDHNTRTTTWVDP 33
| || || | | | | | | | | | | | | | | | |

RESULT 10
US-09-270-767-59345
; Sequence 59345, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59345
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-59345

Query Match          60.6%; Score 134; DB 4; Length 158;
Best Local Similarity 64.9%; Pred. No. 5.9e-11;
Matches 24; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFFIDHNTKTTTWDPRKIP 37
    |||:|||||:|:|:|||||:|:|:|||||
Db 55 GPLPEGWEVRHVTGGRVFIYDHNTRTTQWEDPRLSNP 91

RESULT 11
US-09-270-767-43942
; Sequence 43942, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43942
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43942

Query Match          60.6%; Score 134; DB 4; Length 435;
Best Local Similarity 64.9%; Pred. No. 1.9e-10;
Matches 24; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFFIDHNTKTTTWDPRKIP 37
    |||:|||||:|:|:|||||:|:|:|||||
Db 109 GPLPEGWEVRHVTGGRVFIYDHNTRTTQWEDPRLSNP 145

RESULT 12
US-08-844-312-4
; Sequence 4, Application US/08844312
; Patent No. 5948639
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639el TGF-{SYMBOL 98 \f "Symbol"} Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844.312
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
```

```
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-844-312-4

Query Match          59.3%; Score 131; DB 2; Length 36;
Best Local Similarity 70.0%; Pred. No. 2.7e-11;
Matches 21; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LPKGWEVRHAPNGRPFFIDHNTKTTTWDPR 32
    |||:|||||:|:|:|||||:|:|:|||||
Db 3 LPAGWEQRELPNGRVYVVDHNTKTTTWERP 32

RESULT 13
US-08-630-916A-26
; Sequence 26, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630.916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-916A-26

Query Match          59.3%; Score 131; DB 3; Length 38;
Best Local Similarity 64.7%; Pred. No. 2.9e-11;
Matches 22; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPKGWEVRHAPNGRPFFIDHNTKTTTWDPR 34
    |||:|||||:|:|:|||||:|:|:|||||
Db 1 GPLPSGWEMLTNTARVIFVDHNTKTTTWDPR 34
```

```
RESULT 14
US-08-630-916A-34
; Sequence 34, Application US/08630916A
; Patent No. 601137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Bana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-916A-34

Query Match 59.3%; Score 131; DB 3; Length 38;
Best Local Similarity 70.0%; Pred. No. 2.9e-11;
Matches 21; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LPKGWEVRHAPNGRPPFFIDHNTKTTTWEDP 32
| | | | | | | | | | | | | | | | | |
Db 3 LPAGWEQRELPGRVYVDHNTKTTTWERP 32

RESULT 15
US-08-348-518C-18
; Sequence 18, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
```

```
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEEX: 133521
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Yeast-3
; IMMEDIATE SOURCE:
; CLONE: Rsp5
; US-08-348-518C-18

Query Match 59.3%; Score 131; DB 3; Length 38;
Best Local Similarity 64.7%; Pred. No. 2.9e-11;
Matches 22; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFLPGKWEVRHAPNGRPPFFIDHNTKTTTWEDPRL 34
| | | | | | | | | | | | | | | | | |
Db 1 GPLPSGWEMLTNTARVYFVDHNTKTTTWDDPRL 34

Search completed: October 13, 2005, 14:03:05
Job time : 18.2202 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 81.011 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-5  
Perfect score: 221  
Sequence: 1 GFLPKGVEVRAHPNGRPFDDHNTKTTWEDPRLKIPA 38

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221	100.0	38	14	US-10-185-050-23
2	221	100.0	38	14	US-10-307-956-5
3	221	100.0	733	14	US-10-097-534-16
4	221	100.0	927	14	US-10-097-534-15
5	187	84.6	38	14	US-10-185-050-133
6	187	84.6	38	17	US-10-785-819-169
7	187	84.6	725	14	US-10-185-050-126
8	187	84.6	834	14	US-10-313-955-6
9	187	84.6	854	14	US-10-205-823-277
10	187	84.6	854	14	US-10-287-218-3
11	187	84.6	854	16	US-10-474-291-3

12	187	84.6	854	20	US-11-051-454-277	Sequence 277, App
13	187	84.6	911	14	US-10-205-823-279	Sequence 279, App
14	187	84.6	911	20	US-11-051-454-279	Sequence 279, App
15	187	84.6	923	18	US-10-450-763-31439	Sequence 31439, A
16	187	84.6	975	16	US-10-779-271-3	Sequence 3, Appli
17	187	84.6	995	14	US-10-097-534-9	Sequence 9, Appli
18	187	84.6	995	14	US-10-205-823-275	Sequence 275, App
19	187	84.6	995	20	US-11-051-454-275	Sequence 275, App
20	187	84.6	1071	18	US-10-450-763-42131	Sequence 42131, A
21	149	67.4	38	17	US-10-785-819-158	Sequence 158, App
22	146	66.1	766	14	US-10-313-955-4	Sequence 4, Appli
23	144	65.2	26	14	US-10-185-050-128	Sequence 128, App
24	142	64.3	759	14	US-10-128-714-3162	Sequence 3162, Ap
25	142	64.3	869	14	US-10-128-714-8162	Sequence 8162, Ap
26	136	61.5	38	14	US-10-185-050-22	Sequence 22, Appl
27	136	61.5	38	14	US-10-307-956-8	Sequence 8, Appli
28	136	61.5	832	14	US-10-032-585-7296	Sequence 7296, Ap
29	134	60.6	1082	20	US-11-097-143-16857	Sequence 16857, A
30	131	59.3	38	14	US-10-185-050-26	Sequence 26, Appl
31	131	59.3	38	14	US-10-185-050-34	Sequence 34, Appl
32	131	59.3	38	14	US-10-307-956-9	Sequence 9, Appli
33	131	59.3	38	17	US-10-785-819-163	Sequence 163, App
34	131	59.3	870	14	US-10-097-534-12	Sequence 12, Appl
35	131	59.3	870	16	US-10-723-860-2167	Sequence 2167, Ap
36	131	59.3	906	14	US-10-185-050-48	Sequence 48, Appl
37	129	58.4	38	14	US-10-185-050-134	Sequence 134, App
38	129	58.4	38	17	US-10-785-819-170	Sequence 170, App
39	129	58.4	276	9	US-09-925-300-1527	Sequence 1527, Ap
40	124	56.1	38	14	US-10-307-956-26	Sequence 26, Appl
41	124	56.1	683	14	US-10-185-050-46	Sequence 46, Appl
42	124	56.1	684	14	US-10-097-534-11	Sequence 11, Appl
43	124	56.1	898	15	US-10-188-186-114	Sequence 114, App
44	124	56.1	930	16	US-10-618-408-2	Sequence 2, Appli
45	123	55.7	38	17	US-10-785-819-156	Sequence 156, App

#### ALIGNMENTS

#### RESULT 1

US-10-185-050-23  
; Sequence 23, Application US/10185050  
; Publication No. US20030077577A1  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; Kay, Brian K.  
; Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 233  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/10/185,050  
; APPLICATION NUMBER: US/10/185,050  
; FILING DATE: 28-Jun-2002  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/826,516  
; FILING DATE: 03-Apr-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-208-999



```

; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 133:
US-10-185-050-133

Query Match      84.6%; Score 187; DB 14; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.2e-17;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 37
Db      2 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 37

RESULT 6
US-10-785-819-169
; Sequence 169, Application US/10785819
; Publication No. US20050112552A1
; GENERAL INFORMATION:
; APPLICANT: Herrero J
; APPLICANT: Pirozzi, G.
; APPLICANT: Uveges, A.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING PDZ DOMAINS AND METHODS OF USING SAME
; FILE REFERENCE: 1101-211
; CURRENT APPLICATION NUMBER: US/10/785,819
; CURRENT FILING DATE: 2004-02-23
; PRIOR APPLICATION NUMBER: US/09/723,810
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-785-819-169

Query Match      84.6%; Score 187; DB 17; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.2e-17;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 37
Db      2 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 37

US-10-185-050-126
; Sequence 126, Application US/10185050
; Publication No. US2003007757A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; SEQUENCE CHARACTERISTICS:
; LENGTH: 725 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 126:
US-10-185-050-126

Query Match      84.6%; Score 187; DB 14; Length 725;
Best Local Similarity 88.9%; Pred. No. 1.5e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 37
Db      248 FLPGKGEVVRHAPNGRPPFFIDHNTKTTTWEDPRLKIP 283

RESULT 8
US-10-313-955-6
; Sequence 6, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Caligiuri, Maureen
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313,955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-313-955-6

Query Match      84.6%; Score 187; DB 14; Length 834;
Best Local Similarity 88.9%; Pred. No. 1.8e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2  FLPGGWEVHAPNGRPFIDHNTKTTTWEDPRLKIP 37
Db      357  FLPGGWEVHAPNGRPFIDHNTKTTTWEDPRLKFP 392

RESULT 9
US-10-205-823-277
; Sequence 277, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 277
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-277

Query Match      84.6%; Score 187; DB 14; Length 854;
Best Local Similarity 88.9%; Pred. No. 1.8e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2  FLPGGWEVHAPNGRPFIDHNTKTTTWEDPRLKIP 37
Db      377  FLPGGWEVHAPNGRPFIDHNTKTTTWEDPRLKFP 412

RESULT 10
US-10-287-218-3
; Sequence 3, Application US/10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BUFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; TITLE OF INVENTION: ZEBARJADIAN, Yeganeh
; FILE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 3660046CDI
US-10-287-218-3
```

Query Match	84.6%;	Score 187;	DB 14;	Length 854;
Best Local Similarity	88.9%;	Pred. No. 1.8e-15;		
Matches	32;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	2	FLPKGWEVHAPNGRPFIDHNTKTTTWDPRLKIP	37	
Db	377	FLPPGWEHRIAPNGRPFIDHNTKTTTWDPRLKFP	412	
RESULT 11				
US-10-474-291-3				
Sequence 3,	Application US/10474291			
Publication No.	US20040132043A1			
GENERAL INFORMATION:				
APPLICANT:	AZIMZAI, Yalda; AU-YOUNG, Janice K.			
APPLICANT:	BATRA, Sajeev; BAUGHN, Mariah R.			
APPLICANT:	BECHA, Shanya D.; BOROWSKY, Mark L.			
APPLICANT:	BURFORD, Neil; DING, Li			
APPLICANT:	ELLIOTT, Vicki S.; ENERLING, Brooke M.			
APPLICANT:	GANDHI, Ameena R.; GIETZEN, Kimberly J.			
APPLICANT:	GRIFFIN, Jennifer A.; HAFALIA, April J.A.			
APPLICANT:	HONCHELL, Jennifia D.; LAL, Preeti G.			
APPLICANT:	LEE, Soo Yeun; LIJ, Dyung Aina M.			
APPLICANT:	ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi			
APPLICANT:	REDDY, Roopa M.; SANJANWALA, Madhusudan M.			
APPLICANT:	TANG, Y. Tom; CHAWLA, Narinder K.			
APPLICANT:	WANG, Yu-Mei E.; WARREN, Bridget A.			
APPLICANT:	XU, Yuming; YANG, Junming			
APPLICANT:	YAO, Monique G.; YUE, Henry			
APPLICANT:	ZEBARJADIAN, Yeganeh			
TITLE OF INVENTION:	PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH			
FILE REFERENCE:	PI-0417 USN			
CURRENT APPLICATION NUMBER:	US/10/474,291			
PRIOR FILING DATE:	2003-10-06			
PRIOR APPLICATION NUMBER:	PCT/US02/11152			
PRIOR FILING DATE:	2002-04-05			
PRIOR APPLICATION NUMBER:	US 60/349,705			
PRIOR FILING DATE:	2002-01-15			
PRIOR APPLICATION NUMBER:	US 60/295,263			
PRIOR FILING DATE:	2001-06-01			
PRIOR APPLICATION NUMBER:	US 60/295,340			
PRIOR FILING DATE:	2001-06-01			
PRIOR APPLICATION NUMBER:	US 60/293,727			
PRIOR FILING DATE:	2001-05-25			
PRIOR APPLICATION NUMBER:	US 60/291,846			
PRIOR FILING DATE:	2001-05-18			
PRIOR APPLICATION NUMBER:	US 60/291,662			
PRIOR FILING DATE:	2001-05-16			
PRIOR APPLICATION NUMBER:	US 60/287,228			
PRIOR FILING DATE:	2001-04-27			
PRIOR APPLICATION NUMBER:	US 60/286,820			
PRIOR FILING DATE:	2001-04-26			
PRIOR APPLICATION NUMBER:	US 60/283,294			
PRIOR FILING DATE:	2001-04-11			
Remaining Prior Application data removed - See File Wrapper or PALM.				
NUMBER OF SEQ ID NOS:	42			
SOFTWARE:	PERL Program			
SEQ ID NO 3				
LENGTH:	854			
TYPE:	PRT			
ORGANISM:	Homo sapiens			
NAME/KEY:	misc feature			
OTHER INFORMATION:	Incyte ID No: 3660046CD1			
US-10-474-291-3				
Query Match	84.6%;	Score 187;	DB 16;	Length 854;
Best Local Similarity	88.9%;	Pred. No. 1.8e-15;		
Matches	32;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	2	FLPKGWEVHAPNGRPFIDHNTKTTTWDPRLKIP	37	
Db	377	FLPPGWEHRIAPNGRPFIDHNTKTTTWDPRLKFP	412	

```

; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-279

```

Query Match 84.6%; Score 187; DB 14; Length 911;  
Best Local Similarity 88.9%; Pred. No. 2e-15;  
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy	2	F	L	P	K	G	V	E	R	H	A	N	G	R	P	P	I	D	N	T	K	T	T	W	E	D	P	R	L	K	I	P	37
				:	}																												
D6	434	F	L	P	P	G	W	N	I	A	N	G	R	P	P <td>I</td> <td>D</td> <td>N</td> <td>T</td> <td>K</td> <td>T</td> <td>T</td> <td>W</td> <td>E</td> <td>D</td> <td>P</td> <td>R</td> <td>L</td> <td>K</td> <td>F</td> <td>469</td>	I	D	N	T	K	T	T	W	E	D	P	R	L	K	F	469		

```

RESULT 14
US-11-051-454-279
; Sequence 279, Application US/11051454
; Publication No. US20050191673A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 279
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-051-454-279

```

Query Match 84.6%; Score 187; DB 20; Length 911;  
Best Local Similarity 88.9%; Pred. No. 2e-15;

	Matches	32;	Conservative	1;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	2	FLPKGWVRHAPNGRPPFFIDHNTKTTTWEPRUKIP	37							
Db	434	FLPPGWMRLAPNGRPPFFIDHNTKTTTWEPRUKFP	469							

```

RESULT 15
US-10-450-763-31439
; Sequence 31439, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hvseq. Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP37US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 31439
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (884)..(916)
; OTHER INFORMATION: HECT-domain (ubiquitin-transferase)
; OTHER INFORMATION: eMATRIX, accession number PF00632C, p
; OTHER INFORMATION: 20.86
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (617)..(922)
; OTHER INFORMATION: HECT-domain (ubiquitin-transferase)
; OTHER INFORMATION: Pfam, accession name HECT, E-value=1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(923)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-31439

```

Query Match	84.6%	Score 187;	DB 18;	Length 923;
Best Local Similarity	88.9%	Pred. No. 28-15;		
Matches 32; Conservative		1; Mismatches 3;	Indels 0;	Gaps 0;

Qy	2	F L P K G W E V R H A P N G R P F F I D H N T K T T T W E D P R L K I P 37
	:	
Db	446	F L P P G W E M R I A P N G R P F F I D H N T K T T T W E D P R L K F P 481

Search completed: October 13, 2005, 14:50:01  
Job time : 82.011 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:36:01 ; Search time 15.8681 Seconds  
(without alignments)  
230.414 Million cell updates/sec

Title: US-09-385-918-5

Perfect score: 221

Sequence: 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIPA 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	84.6	820	2 T46412	ubiquitin-protein
2	146	66.1	766	1 S66562	ubiquitin-protein
3	146	66.1	767	2 T37545	ubiquitin-protein
4	143	64.7	815	2 T49744	probable ubiquitin
5	136	61.5	786	2 T39585	ubiquitin-protein
6	136	61.5	809	1 S43217	ubiquitin-protein
7	131	59.3	1256	2 JF0209	brain-specific ang
8	125	56.6	708	2 I83196	NEDD-4 ORF - mouse
9	125	56.6	887	2 S70642	ubiquitin ligase N
10	123	55.7	1012	2 T23160	hypothetical prote
11	121.5	55.0	616	2 T46292	hypothetical prote
12	121	54.8	1171	2 T42372	probable guanylate
13	118	53.4	457	2 G89632	protein F13E6.4 [i
14	115	52.0	671	2 T37900	probable ubiquitin
15	115	52.0	1277	2 T14152	synaptic scaffoldi
16	113	51.1	472	2 B56954	yes-associated pro
17	107	48.4	448	2 I50730	yes-associated pro
18	104.5	47.3	1616	2 T47801	hypothetical prote
19	99	44.8	454	2 A56954	yes-associated pro
20	95	43.0	383	2 JC7507	45K WW domain-cont
21	95	43.0	386	2 JC7508	45K WW domain-cont
22	87	39.4	691	2 T33637	hypothetical prote
23	82	37.1	870	2 A41130	dystrophin homolog
24	82	37.1	3660	1 S02041	dystrophin, muscle
25	82	37.1	3678	2 S28916	dystrophin - mouse
26	82	37.1	3685	1 A27605	dystrophin, muscle
27	73.5	33.3	67	2 S64713	formin binding pro
28	72	32.6	241	2 T25886	hypothetical prote
29	71.5	32.4	889	2 T20274	hypothetical prote

RESULT 1

T46412

ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)

N:Alternate names: hypothetical protein DKFZp434P2422.1

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C:Accession: T46412

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223034

A:Accession: T46412

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-820 <AAA>

A:Cross-references: UNIPROT:Q9NT88; EMBL:AL137469

A:Experimental source: adult testis; clone DKFZp434P2422

C:Genetics:

A:Gene: GDB:NEDD4

A:Note: DKFZp434P2422.1

C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

C:Keywords: ligase

F:142-179/Domain: WW repeat homology <WWR1>

F:342-379/Domain: WW repeat homology <WWR2>

F:393-430/Domain: WW repeat homology <WWR3>

F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 84.6%; Score 187; DB 2; Length 820;  
Best Local Similarity 88.9%; Pred. No. 2.1e-16;  
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FLPLKGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIP 37

DB 343 FLPLFGWEVRHAPNGRPFIDHNTKTTTWEDPRLKIP 378

RESULT 2

S66562

ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: E6-AP-like protein ubiquitin ligase

C:Species: Schizosaccharomyces pombe

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S66562; T45159

R:Nefsky, B.; Beach, D.

EMBO J. 15, 1301-1312, 1996

A:Title: Pbl1 acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25

A:Reference number: S66562; MUID:96205868; PMID:8635463

A:Accession: S66562

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-766 <NEF>

A:Cross-references: UNIPROT:Q92462; GB:U66716; NID:gl519443; PIDN:AAB07514.1; PID:gl5194

R:Nefsky, B.S.; Beach, D.

submitted to the EMBL Data Library, August 1996  
A:Description: Pub1 acts as an E6-Ap-like protein ubiquitin ligase in the degradation of  
A:Reference number: Z22935  
A:Accession: T45159  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-766 <NE2>  
A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1  
C:Genetics:  
A:Gene: Pub1  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C:Keywords: ligase  
F:205-242/Domain: WW repeat homology <WW1>  
F:288-325/Domain: WW repeat homology <WW2>  
F:345-382/Domain: WW repeat homology <WW3>  
Query Match 66.1%; Score 146; DB 1; Length 766;  
Best Local Similarity 69.7%; Pred. No. 4.6e-11;  
Matches 23; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPR 33  
Db 288 GELPPGWEQRYTPGRPYFVDHNTTRTTTWVDP 320  
RESULT 3  
T37545  
ubiquitin-protein ligase (EC 6.3.2.19) publ [validated] - fission yeast (Schizosaccharom  
N:Alternate names: ubiquitin ligase Pub1  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T37545; T48655  
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
Submitted to the EMBL Data Library, September 1997  
A:Reference number: Z21722  
A:Accession: T37545  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-767 <MCL>  
A:Cross-references: UNIPROT:Q92462; EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GNO00666; SPDB:SF  
A:Experimental source: strain 972h-; cosmid c117  
R:Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.  
Mol. Gen. Genet. 254, 520-528, 1997  
A:Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning publ ub  
A:Reference number: Z07985; MUID:97340937; PMID:9197411  
A:Accession: T48655  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-767 <SAL>  
A:Cross-references: EMBL:U62795; NID:q2262192; PIDN:AAB63350.1; PID:q2262193  
A:Experimental source: strain J27  
C:Genetics:  
A:Gene: publ; SPDB:SPAC11G7.02  
A:Map position: 1  
A:Introns: 6/2; 14/1; 62/2  
C:Function: <UBI>  
A:Description: EC 6.3.2.19 [validated, MUID:96205868]  
A:Function: <CVC>  
A:Description: involved in of the mitotic activating tyrosine phosphatase cdc25 [validat  
C:Function: <TOL>  
A:Description: required for low pH-tolerance [validated, MUID:97340937]  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C:Keywords: cell cycle control; ligase  
F:205-242/Domain: WW repeat homology <WW1>  
F:288-325/Domain: WW repeat homology <WW2>  
F:345-382/Domain: WW repeat homology <WW3>  
Query Match 66.1%; Score 146; DB 2; Length 767;  
Best Local Similarity 69.7%; Pred. No. 4.6e-11;  
Matches 23; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPR 33

Db 288 GELPPGWEQRYTPGRPYFVDHNTTRTTTWVDP 320  
RESULT 4  
T49744  
probable ubiquitin-protein ligase [imported] - Neurospora crassa  
N:Alternate names: protein B24B19.160  
C:Species: Neurospora crassa  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C:Accession: T49744  
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49744  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-815 <SCH>  
A:Cross-references: EMBL:AL356192; GSPDB:GNO0116; NCSP:B24B19.160  
A:Experimental source: BAC clone B24B19; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B24B19.160  
A:Map position: 6  
A:Introns: 11/1; 24/1; 59/2; 110/1; 783/2  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
F:239-276/Domain: WW repeat homology <WWR1>  
F:334-371/Domain: WW repeat homology <WWR2>  
F:393-430/Domain: WW repeat homology <WWR3>  
Query Match 64.7%; Score 143; DB 2; Length 815;  
Best Local Similarity 69.7%; Pred. No. 1.2e-10;  
Matches 23; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 GFLPKGWEVRHAPNGRPFIDHNTKTTTWEDPR 33  
Db 334 GELPAGWEQRTPEGRPYFVDHNTTRTTTWVDP 366  
RESULT 5  
T39585  
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T39585  
R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21865  
A:Accession: T39585  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-786 <VOL>  
A:Cross-references: UNIPROT:O14326; EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GNO0067; SPDB:SP  
A:Experimental source: strain 972h-; cosmid c1859  
C:Genetics:  
A:Gene: SPDB:SPBC16E9.11c  
A:Map position: 2  
A:Introns: 60/2  
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
F:236-273/Domain: WW repeat homology <WWR1>  
F:306-343/Domain: WW repeat homology <WWR2>  
F:364-401/Domain: WW repeat homology <WWR3>  
Query Match 61.5%; Score 136; DB 2; Length 786;  
Best Local Similarity 67.7%; Pred. No. 9.6e-10;  
Matches 21; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
Qy 3 LPKGWEVRHAPNGRPFIDHNTKTTTWEDPR 33  
Db 308 LPFGWEMRYTDTGRPYFVDHNTTRTTTWVDP 338  
RESULT 6  
S43217  
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)



Query Match 61.5%; Score 136; DB 1; Length 809;  
 Best Local Similarity 66.7%; Pred. No. 9.9e-10;  
 Matches 22; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GELPKGWEVHAPNGRPFDFIDHNTKTTTWEDPR 33  
 | | | | | | | | | | | | | | | | | | | |  
 DB 331 GELPSGWEQRTPEGRAYFVDHNTKTTTWEDPR 363  
 | | | | | | | | | | | | | | | | | | | |

RESULT 7  
 JE0209  
 brain-specific angiogenesis inhibitor-associated protein 1 - human  
 N/Alternate names: BAI1-associated protein 1; BAP1 [misnomer]  
 C/Species: Homo sapiens (man)  
 C/Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
 C/Accession: JE0209  
 R/Shiratsuchi, T.; Futamura, M.; Oda, K.; Nishimori, H.; Nakamura, Y.; Tokino, T.  
 Biochem. Biophys. Res. Commun. 247, 597-604, 1998  
 A/Title: Cloning and characterization of BAI1-associated protein 1: A PDZ domain-containing  
 A/Reference number: JE0209; MUID:98321173; PMID:9647739  
 A/Accession: JE0209  
 A/Molecule type: mRNA  
 A/Residues: 1-1256 <SH1>  
 A/Cross-references: UNIPROT:O75085; UNIPROT:Q96QZ7; GB:AB010894; NID:G3370997  
 A/Experimental source: brain  
 C/Genetics:  
 A/Gene: GDB:BA1AP1; BAP1  
 A/Cross-references: GDB:9864783  
 A/Map position: 3p14.1-3p14.1  
 F/300-337/Domain: WW repeat homology <WW1>  
 F/339-396/Domain: WW repeat homology <WW2>

Query Match 59.3%; Score 131; DB 2; Length 1256;  
 Best Local Similarity 63.6%; Pred. No. 7.5e-09;



RESULT 15  
T14152  
synaptic scaffolding protein S-SCAM - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 15-Mar-2004  
C:Accession: T14152  
R: Hirao, K.; Hata, Y.; Ide, N.; Takeuchi, M.; Irie, M.; Yao, I.; Deguchi, M.; Toyoda, A.  
J. Biol. Chem. 273, 21105-21110, 1998  
A: Title: A novel multiple PDZ domain-containing molecule interacting with N-methyl-D-aspartate receptor  
A: Reference number: Z17889; MUID: 98361985; PMID: 9694864  
A: Accession: T14152  
A: Status: preliminary; translated from GB/EMBL/DDBJ  
A: Molecule type: mRNA  
A: Residues: 1-1277 <HIR>  
A: Cross-references: EMBL:AF034863; NID:G3411052; PID:G3411053; PIDN:AAC31124.1  
C: Genetics:  
A: Gene: S-SCAM  
C: Function:  
A: Description: may assemble receptors and cell adhesion proteins at synaptic junctions  
F: 302-339/Domain: WW repeat homology <WW1>  
F: 348-385/Domain: WW repeat homology <WW2>

Query Match 52.0%; Score 115; DB 2; Length 1277;  
Best Local Similarity 59.4%; Pred. No. 9.5e-07;  
Matches 19; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 3 LPKGWEVRHAPNGRPFIDHNTKTTTWDPR 34  
||| : : : : : : : : : : : : : : : : : :  
Db 304 LPDNWEMAYTEKGEVYFIDHNTKTTSWLDPRL 335

Search completed: October 13, 2005, 15:11:15  
Job time : 15.8681 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 72.381 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-5  
Perfect score: 221  
Sequence: 1 GFLPKGEVHRAPNGRPFIDHNTKTTTWDPRLKIPA 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221	100.0	1000	1 NED4_HUMAN	P46934 homo sapien
2	215	97.3	455	2 Q9N134	Q9N134 oryctolagus
3	187	84.6	571	2 Q8QG2	Q8QG2 gallus gall
4	187	84.6	820	2 Q9NT88	Q9NT88 homo sapien
5	187	84.6	835	2 Q8BRT9	Q8BRT9 mus musculus
6	187	84.6	854	2 Q8H2W4	Q8H2W4 homo sapien
7	187	84.6	855	2 Q8BW58	Q8BW58 homo sapien
8	187	84.6	855	2 Q99PK2	Q99PK2 mus musculus
9	187	84.6	855	2 Q8CF10	Q8CF10 mus musculus
10	187	84.6	875	2 Q641N8	Q641N8 mus musculus
11	187	84.6	911	2 Q8NSA7	Q8NSA7 homo sapien
12	187	84.6	947	2 Q7Z5F1	Q7Z5F1 homo sapien
13	187	84.6	955	2 Q86PU5	Q86PU5 homo sapien
14	187	84.6	967	2 Q7Z5F2	Q7Z5F2 homo sapien
15	187	84.6	971	2 Q42573	Q42573 xenopus lae
16	187	84.6	971	2 Q6GMD5	Q6GMD5 xenopus lae
17	187	84.6	975	2 Q7Z5N3	Q7Z5N3 homo sapien
18	187	84.6	995	2 Q43165	Q43165 homo sapien
19	171.5	77.6	970	2 Q6DIR6	Q6DIR6 xenopus tro
20	146	66.1	767	1 PUB1_SCHPO	Q92462 schizosacch
21	143	64.7	787	2 Q6M906	Q6M906 neurospora
22	143	64.7	806	2 Q7RV01	Q7RV01 neurospora
23	142	64.3	854	2 Q6C5H0	Q6C5H0 varrowia li
24	141	63.8	917	2 Q7QE76	Q7QE76 anopheles g
25	138	62.4	724	2 Q8BKW4	Q8BKW4 caenorhabdi
26	138	62.4	817	2 Q7SA12	Q7SA12 ashbya goss
27	136	61.5	786	1 PUB3_SCHPO	O43326 schizosacch
28	136	61.5	809	1 RP5_YEAST	P39940 saccharomyc
29	136	61.5	822	2 Q6FN71	Q6FN71 candida gla
30	135	61.1	781	2 Q6BT41	Q6BT41 debaryomyce
31	135	61.1	819	2 Q6CNC7	Q6CNC7 kluyveromyce

RESULT 1  
NED4\_HUMAN  
ID NED4\_HUMAN STANDARD; PRT; 1000 AA.  
AC P46934;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE E3 ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).  
OS Name=NEDD4; Synonyms=KIAA0093;  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP Hinz U.;  
RL Unpublished observations (JUN-2003).  
RN (2)  
RP SEQUENCE OF 74-1000 FROM N.A.  
RX MEDLINE=95308325; PubMed=7788527;  
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,  
Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. III.  
The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by  
analysis of cDNA clones from human cell line KG-1.";  
DNA Res. 2:37-43(1995).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
then directly transfers the ubiquitin to targeted substrates (By  
similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Interacts with UBE2D2. Binds SCNN1A, SCNN1B and SCNN1G  
(By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
thiolester formation.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
domain.  
CC -!- SIMILARITY: Contains 4 WW domains.  
CC -!- CAUTION: The sequence of the N-terminus was deduced from the  
genomic sequence.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; AC039057; -; NOT ANNOTATED\_CDS.  
EMBL; D42055; BAA07655.1; -;



```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 C2 domain.
DR EMBL; AL137469; CAB70754.1; -.
DR HSSP; Q62940; 115H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00403; WNDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00004; C2 DOMAIN 2; 1.
DR PROSITE; PS0237; HECT; 1_2; 1.
DR PROSITE; PS01159; WW_DOMAIN 1; 3.
DR PROSITE; PS00020; WW_DOMAIN 2; 3.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 820 AA; 95282 MW; 0FDB34B29B3F4123 CRC64;

Query Match 84.6%; Score 187; DB 2; Length 820;
Best Local Similarity 88.9%; Pred. No. 2.6e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FLPGWEVRHAPNGRPFIDHTKTTTWEEDRLKIP 37
Db 343 FLPGWEVRHAPNGRPFIDHTKTTTWEEDRLKFP 378

RESULT 5
Q8BRT9 PRELIMINARY; PRT; 835 AA.
ID Q8BRT9
AC Q8BRT9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:A73001F13 product:neural cell expressed,
DE developmentally down-regulated gene 4b, full insert sequence.
GN Name=Nedd41;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";

```

```

RL Nature 409:695-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto K., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK042621; BACJ1307.1; -.
DR HSSP; Q62940; 115H.
DR MGD; MGI:1933754; Nedd41.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WNDOMAIN.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN 1; 4.
DR PROSITE; PS00020; WW_DOMAIN 2; 4.
SQ SEQUENCE 835 AA; 96410 MW; 81C7DD62722DDA9B CRC64;

Query Match 84.6%; Score 187; DB 2; Length 835;
Best Local Similarity 88.9%; Pred. No. 2.6e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FLPGWEVRHAPNGRPFIDHTKTTTWEEDRLKIP 37
Db 358 FLPGWEVRHAPNGRPFIDHTKTTTWEEDRLKFP 393

```

```
RESULT 6
Q9H2W4 ID Q9H2W4 PRELIMINARY; PRT; 854 AA.
AC Q9H2W4
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NEDD4L (Ubiquitin ligase NEDD4Lb) (Ubiquitin ligase NEDD4f).
GN Name=NEDD4L; Synonyms=NEDD4L;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21828840; PubMed=11840194; DOI=10.1038/sj.ejhg.5200747;
RA Chen H., Ross C.A., Wang N., Huo Y., MacKinnon D.F., Potash J.B.,
RA Simpson S.G., McMahon P.J., DePaulo J.R. Jr., McInnis M.G.;
RT "NEDD4L on human chromosome 18q21 has multiple forms of transcripts
and is a homologue of the mouse Nedd4-2 gene.";
RL Eur. J. Hum. Genet. 9:922-930(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009;
RA Qi H., Grenier J., Fournier A., Labrie C.;
RT "Androgens differentially regulate the expression of NEDD4L
transcripts in LNCaP human prostate cancer cells.";
RL Mol. Cell. Endocrinol. 210:51-62(2003).
DR EMBL; AF210730; AAG43524.1; -
DR EMBL; AF385931; AAM46208.1; -
DR EMBL; AY112983; AAM76728.1; -
DR HSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Reps_WWP.
DR Pfam; PF00632; HECT; 1.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00456; WW; 4.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
KW Ligase.
SQ SEQUENCE 854 AA; 98180 MW; 00C74E1661F52E7F CRC64;
Query Match 84.6%; Score 187; DB 2; Length 854;
Best Local Similarity 88.9%; Pred. No. 2.6e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FLPGKGVHRHAPNGRPFIDHNTKTTTWEDPRLKIP 37
Db 377 FLPPGWMRIAPNGRPFIDHNTKTTTWEDPRLKFP 412

RESULT 7
Q9BW58 ID Q9BW58 PRELIMINARY; PRT; 855 AA.
AC Q9BW58; Q8WU09;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NEDD4L protein (Hypothetical protein) (Fragment).
GN Name=NEDD4L;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009;
RA Qi H., Grenier J., Fournier A., Labrie C.;
RT "Androgens differentially regulate the expression of NEDD4L
transcripts in LNCaP human prostate cancer cells.";
RL Mol. Cell. Endocrinol. 210:51-62(2003).
DR EMBL; AF210730; AAG43524.1; -
DR EMBL; AF385931; AAM46208.1; -
DR EMBL; AY112983; AAM76728.1; -
DR HSP; Q62940; I15H.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Reps_WWP.
DR Pfam; PF00632; HECT; 1.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00456; WW; 4.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
KW Hypothetical protein.
SQ SEQUENCE 855 AA; 98604 MW; 3CF4A66996F033EA CRC64;
Query Match 84.6%; Score 187; DB 2; Length 855;
Best Local Similarity 88.9%; Pred. No. 2.6e-15;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FLPGKGVHRHAPNGRPFIDHNTKTTTWEDPRLKIP 37
Db 378 FLPPGWMRIAPNGRPFIDHNTKTTTWEDPRLKFP 413

RESULT 8
Q99PK2 ID Q99PK2 PRELIMINARY; PRT; 855 AA.
AC Q99PK2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```



DE Ubiquitin-protein ligase Nedda4-2.  
GN Name=Nedd41;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ON NCBI\_TaxID=10090;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6 Ncr1 BR;  
RX MEDLINE=21067027; PubMed=11149908;  
RA Kamylna E., Debonneville C., Bens M., Vandewalle A., Staub O.;  
RT "A novel mouse Nedda4 protein suppresses the activity of the epithelial  
RT Na+ channel.";  
RL FASEB J. 15:204-214(2001).  
DR EMBL; AF277232; AAK00809.1; -.  
DR HSSP; Q62940; 115H.  
DR MGD; MGI:1933754; Nedda4.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00119; HECTc; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 4.  
KW Ligase.  
SQ SEQUENCE 855 AA; 98258 MW; 3BE1566E2C3BA20E CRC64;  
  
Query Match 84.6%; Score 187; DB 2; Length 855;  
Best Local Similarity 88.9%; Pred. No. 2.6e-15;  
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 2 FLPGKGVHRHAPNGRPFIDHNTKTTTWTWEDPRLKIP 37  
Db 378 FLPGWEMRIAPNGRPFIDHNTKTTTWTWEDPRLKFP 413  
  
RESULT 9  
Q8CF10  
ID Q8CF10 PRELIMINARY; PRT; 855 AA.  
AC Q8CF10;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Neural cell expressed, developmentally down-regulated gene 4-  
DE like.  
GN Name=Nedd41;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ON NCBI\_TaxID=10090;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Kidney, and Mammary tumor. C3;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
RA Strausberg R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC039746; AAH39746.1; -.  
DR EMBL; BC071210; AAH71210.1; -.  
DR HSSP; Q62940; 115H.  
DR MGD; MGI:1933754; Nedda4.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00119; HECTc; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 4.  
SQ SEQUENCE 855 AA; 98465 MW; 96C452B442855895 CRC64;  
  
Query Match 84.6%; Score 187; DB 2; Length 855;  
Best Local Similarity 88.9%; Pred. No. 2.6e-15;  
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 2 FLPGKGVHRHAPNGRPFIDHNTKTTTWTWEDPRLKIP 37  
Db 378 FLPGWEMRIAPNGRPFIDHNTKTTTWTWEDPRLKFP 413  
  
RESULT 10  
Q641N8  
ID Q641N8 PRELIMINARY; PRT; 875 AA.  
AC Q641N8;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Nedda4 protein (fragment).  
GN Name=Nedd41;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ON NCBI\_TaxID=10090;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

HSSP; Qc2940; IISH.  
GO; GO:0005622; C:intracellular; IEA.  
GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 3.  
DR PRINTS; PR00360; C2DOMAIN.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS00237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 3.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 3.  
DR SEQUENCE 911 AA; 104921 MW; CE04AAED677AA506 CRC64;  
  
Query Match 84.6%; Score 187; DB 2; Length 911;  
Best Local Similarity 88.9%; Pred. No. 2.8e-15;  
Matches 32; Conservative 1; Mismatches 3; Indels 0;  
  
QY 2 FLPPGWEVRHAPNGRPFFIDHTKTTTWTEDPRLKIP 37  
||| ||| : ||| ||| ||| ||| ||| ||| ||| |||  
DB 434 FLPPGWEVRHAPNGRPFFIDHTKTTTWTEDPRLKFP 469  
  
RESULT 12  
Q7Z5F1 PRELIMINARY; PRT; 947 AA.  
AC AC Q7Z5F1;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DI 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Ubiquitin ligase NEDD4h.  
GN Name=NEDD4L;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22380412; PubMed=14615060; DOI=10.1016/j.mce.2003.08.009  
RA Qi H., Grenier J., Fournier A., Labrie C.;  
RT "Androgens differentially regulate the expression of NEDD4L transcripts in LNCap human prostate cancer cells.";  
RT Mol. Cell. Endocrinol. 210:51-62(2003).  
RC -I- SIMILARITY: Contains 1 C2 domain.  
DR EMBL; AY112985; AAM76730.1; -.  
DR HSSP; Qc2940; IISH.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0018874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR000569; HECT.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS00237; HECT; 1.

**DR PROSITE**; PS01159; WW\_DOMAIN\_1; 4.  
**DR PROSITE**; PSS0020; WW\_DOMAIN\_2; 4.  
**KW Ligase.**

SQ SEQUENCE	947 AA;	109429 MW;	95CBFDEDCB96639 CRC64;
-------------	---------	------------	------------------------

  

<b>Query Match</b>			
Best Local Similarity	84.6%;	Score 187;	DB 2; Length 947;
Matches	32;	Conservative	1; Mismatches 3; Indels 0; Gaps 0;

```

      2 FLPKGEVVRHAPNGRPFIDHTKTITTTTWEPRLKIP 37
      |||||:|||||:|||||:|||||:|||||:|||||:|||||
    470 FLPPGWMRIAPNGRPFFDHNTHNTTKTTTWEDRLKFP 505
  
```

**RESULT 13**

ID	Q96PU5 PRELIMINARY;	PRT;	955 AA.
AC Q96PU5,			
DT 01-DEC-2001 (TREMBLrel. 19, Created)			
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE NEDD4-like ubiquitin ligase 3.			
GN Name=NEDL3;			
OS Homo sapiens (Human)			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxId=9606;			
RN [1]			
RN NAME=NEDDL3;			
RN SEQUENCE FROM N.A.			
RP MEDLINE=20501262; PubMed=110461148;			
RX DOI=10.1128/MCB.20.22.8526-8535.2000;			
RA Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G., Ingham R., Ernberg I., Pawson T.; "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3 protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases." Mol. Cell. Biol. 20:8526-8535(2000).			
RN [2]			
RN SEQUENCE FROM N.A.			
RP MEDLINE=21269431; PubMed=11244092; DOI=10.1074/jbc.C000906200;			
RX Harvey F.K., Dinudom A., Cook I.D., Kumar S.; "The Nedda4-like protein KIAA0439 is a potential regulator of the epithelial sodium channel." J. Biol. Chem. 276:8597-8601(2001).			
RN [3]			
RN SEQUENCE FROM N.A.			
RP Okamoto Y., Miyazaki K., Sakamoto M., Kato C., Nakagawara A.; Submitted (SEP-2001) to the ENBL/GenBank/DBJ databases.			
CC !- SIMILARITY: Contains 1 C2 domain.			
EMBL; AB071179; BAB69424.1; .			
DR HSSP; Q62940; 1I5H.			
DR GO; GO:0005622; C:intracellular; IC.			
DR GO; GO:0005515; F:protein binding; IP.			
DR GO; GO:0017080; F:sodium channel regulator activity; IDA.			
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.			
DR GO; GO:0007588; P:excretion; NAS.			
DR GO; GO:0045807; P:positive regulation of endocytosis; NAS.			
DR GO; GO:0016567; P:protein ubiquitination; NAS.			
DR GO; GO:0042176; P:regulation of protein catabolism; NAS.			
DR GO; GO:0010038; P:response to metal ion; IDA.			
DR GO; GO:0006883; P:sodium ion homeostasis; NAS.			
DR GO; GO:0006814; P:sodium ion transport; NAS.			
DR GO; GO:0030104; P:water homeostasis; NAS.			
DR InterPro; IPR000008; C2.			
DR InterPro; IPR008973; C2_CalB.			
DR InterPro; IPR000569; HECT.			
DR InterPro; IPR002349; WW.			
DR InterPro; IPR001202; WW_Rsp5_WWP.			
DR Pfam; PF00168; C2; 1.			
DR Pfam; PF00632; HECT; 1.			
DR PRINTS; PR00360; C2DOMAIN.			
DR PRINTS; PR00403; WWDOMAIN.			
DR SMART; SM00239; C2; 1.			

Dr	SMART; SM00119; HECTC; 1.
Dr	SMART; SM00456; WW; 4.
Dr	PROSITE; PS00499; C2 DOMAIN_1; UNKNOWN_1.
Dr	PROSITE; PSS0004; C2 DOMAIN_2; 1.
Dr	PROSITE; PS0237; HECT; 1.
Dr	PROSITE; PS01159; WW DOMAIN_1; 4.
Dr	PROSITE; PSS0020; WW DOMAIN_2; 4.
Kw	Ligase.
Sq	SEQUENCE 955 AA; 110021 MW; ABB278A37F6A6B5 CRC64;

  

<b>Query Match</b>			
Best Local Similarity	84.6%;	Score 187;	DB 2; Length 955;
Matches	32;	Conservative	1; Mismatches 3; Indels 0; Gaps 0;

```

      2 FLPKGEVVRHAPNGRPFIDHTKTITTTTWEPRLKIP 37
      |||||:|||||:|||||:|||||:|||||:|||||:|||||
    478 FLPPGWMRIAPNGRPFFDHNTHNTTKTTTWEDRLKFP 513
  
```

**RESULT 14**

ID	Q7ZSF2 PRELIMINARY;	PRT;	967 AA.
AC Q7ZSF2,			
DT 01-OCT-2003 (TREMBLrel. 25, Created)			
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE Ubiquitin ligase NECD4g.			
GN Name=NECD4L;			
OS Homo sapiens (Human)			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxId=9606;			
RN [1]			
RN NAME=NECD4L;			

Search completed: October 13, 2005, 15:09:23  
Job time : 76.381 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 13.5165 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-15  
Perfect score: 30  
Sequence: 1 XPPPY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A: Geneseq\_16Dec04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	5	ADA50768	Viral inh
2	28	93.3	6	AAB83031	Human Sma
3	28	93.3	8	ADA50788	Viral inh
4	28	93.3	9	ADK84421	Human inh
5	28	93.3	9	ADK88045	Human 191
6	28	93.3	9	ADK88545	Human 191
7	28	93.3	9	ADK88044	Human 191
8	28	93.3	9	ADK84919	Human 191
9	28	93.3	9	ADK84968	Human 191
10	28	93.3	9	ADK88384	Human 191
11	28	93.3	9	ADK84460	Human 191
12	28	93.3	9	ADK86562	Human 191
13	28	93.3	9	ADK87588	Human 191
14	28	93.3	9	ADK83902	Human 191
15	28	93.3	9	ADK88578	Human 191
16	28	93.3	9	ADK88349	Human 191
17	28	93.3	9	ADK85989	Human 191
18	28	93.3	9	ADK86505	Human 191
19	28	93.3	9	ADK85516	Human 191
20	28	93.3	9	ADK87020	Human 191
21	28	93.3	10	AAR98323	WBP-1 PY
22	28	93.3	10	ADK83656	Human 191
23	28	93.3	10	ADK84646	Human 191
24	28	93.3	10	ADK86759	Human 191
25	28	93.3	10	ADK86754	Human 191

26	28	93.3	10	ADK85170	Human 191
27	28	93.3	10	ADK85755	Human 191
28	28	93.3	10	ADK89547	Human 191
29	28	93.3	10	ADK83605	Human 191
30	28	93.3	10	ADK86224	Human 191
31	28	93.3	10	ADK88674	Human 191
32	28	93.3	10	ADK85778	Human 191
33	28	93.3	10	ADK84124	Human 191
34	28	93.3	10	ADK86828	Human 191
35	28	93.3	10	ADK89444	Human 191
36	28	93.3	10	ADK89120	Human 191
37	28	93.3	10	ADK89613	Human 191
38	28	93.3	10	ADK89614	Human 191
39	28	93.3	10	ADK84704	Human 191
40	28	93.3	10	ADK89376	Human 191
41	28	93.3	10	ADK83562	Human 191
42	28	93.3	12	AAW38128	PPPY mot
43	28	93.3	12	ADB49367	PPPY mot
44	28	93.3	13	AAW38123	PPPY mot
45	28	93.3	13	ADB49362	PPPY mot

ALIGNMENTS

RESULT 1  
ADA50768  
ID ADA50768 standard; peptide; 5 AA.  
XX AC ADA50768;  
XX AC ADA50768;  
DT 20-NOV-2003 (first entry)  
XX Viral inhibitor peptide SEQ ID NO:5.  
XX PPXY motif; type I WW-domain; Nedd4; virucide; viral budding;  
KW virus propagation; viral infection; hepatitis B virus;  
KW human herpes virus 1.  
XX Unidentified.  
XX WO2003015714-A2.  
XX 27-FEB-2003.  
XX 21-AUG-2002; 2002WO-US026681.  
XX 21-AUG-2001; 2001US-0313883P.  
XX (MYRI-) MYRIAD GENETICS INC.  
XX Morham S, Zavitz K, Hobden A;  
XX WPI; 2003-300559/29.  
XX Novel peptide capable of binding to type I WW-domain of Nedd4 protein and useful for treating viral infections, comprises contiguous amino acid sequence of a viral protein encompassing PPXY motif.  
XX Disclosure; Page 12; 75pp; English.  
XX The invention relates to a novel isolated peptide comprising a contiguous amino acid sequence of 8-30 amino acid residues of a viral protein, where the contiguous amino acid sequence encompasses the PPXY motif of the viral protein and the peptide is capable of binding a type I WW-domain of the Nedd4 protein. A peptide of the invention has virucide activity, and acts as an inhibitor of viral budding and virus propagation. The peptides are useful in the manufacture of a medicament useful in the treatment of viral infections caused by a virus selected from hepatitis B virus and human herpes virus 1. The peptides are also useful for inhibiting viral budding from virus-infected cells and thus inhibiting virus propagation in the cells. The present sequence represents a peptide of the invention.

*Handwritten signature*



XX DE Human 191P4D12(b) peptide fragment #1122.  
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
XX KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX OS Homo sapiens.  
XX PN WO2004016799-A2.  
XX PD 26-FEB-2004.  
XX PF 23-APR-2003; 2003WO-US013013.  
XX PR 16-AUG-2002; 2002US-0404306P.  
XX PR 01-NOV-2002; 2002US-0423290P.  
XX PA (AGEN-) AGENSYS INC.  
XX PI Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX PD 26-FEB-2004.  
XX PF 23-APR-2003; 2003WO-US013013.  
XX PR 16-AUG-2002; 2002US-0404306P.  
XX PR 01-NOV-2002; 2002US-0423290P.  
XX PA (AGEN-) AGENSYS INC.  
XX PI Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX PD New composition comprising 191P4D12(b) proteins and polynucleotides,  
XX PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
XX PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
XX PT humoral or cellular immune response.  
XX PS Claim 1; Page 154; 443pp; English.  
XX CC The present invention relates to novel compositions comprising peptides  
XX CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
XX CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
XX CC expressed in a number of cancers and so the compositions of the invention  
XX CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
XX CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
XX CC uterus or cervix and in eliciting a humoral or cellular immune response.  
XX CC To isolate genes that are overexpressed in prostate cancer, the  
XX CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
XX CC derived from prostate cancer tissues.  
XX SQ Sequence 9 AA;  
XX Query Match 93.3%; Score 28; DB 8; Length 9;  
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 2 PPPXY 6  
Db 3 PPSY 7  
RESULT 5  
ADK88045  
ID ADK88045 standard; peptide; 9 AA.  
XX AC ADK88045;  
XX XX Homo sapiens.  
XX DT 20-MAY-2004 (first entry)  
XX DE Human 191P4D12(b) peptide fragment #4746.  
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
XX KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
XX KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
XX KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
XX KW Suppression Subtractive Hybridisation; SSH.  
XX OS Homo sapiens.

PN WO2004016799-A2.  
XX PD 26-FEB-2004.  
XX PF 23-APR-2003; 2003WO-US013013.  
XX PR 16-AUG-2002; 2002US-0404306P.  
XX PR 01-NOV-2002; 2002US-0423290P.  
XX PA (AGEN-) AGENSYS INC.  
XX PI Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX PD New composition comprising 191P4D12(b) proteins and polynucleotides,  
XX PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
XX PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
XX PT humoral or cellular immune response.  
XX PS Claim 1; Page 199; 443pp; English.  
XX CC The present invention relates to novel compositions comprising peptides  
XX CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
XX CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
XX CC expressed in a number of cancers and so the compositions of the invention  
XX CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
XX CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
XX CC uterus or cervix and in eliciting a humoral or cellular immune response.  
XX CC To isolate genes that are overexpressed in prostate cancer, the  
XX CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
XX CC derived from prostate cancer tissues.  
XX SQ Sequence 9 AA;  
XX Query Match 93.3%; Score 28; DB 8; Length 9;  
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 2 PPPXY 6  
Db 1 PPSY 5  
RESULT 6  
ADK88545  
ID ADK88545 standard; peptide; 9 AA.  
XX AC ADK88545;  
XX XX Homo sapiens.  
XX DT 20-MAY-2004 (first entry)  
XX DE Human 191P4D12(b) peptide fragment #5246.  
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
XX KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
XX KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
XX KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
XX KW Suppression Subtractive Hybridisation; SSH.  
XX OS Homo sapiens.  
XX PN WO2004016799-A2.  
XX PD 26-FEB-2004.  
XX PF 23-APR-2003; 2003WO-US013013.  
XX PR 16-AUG-2002; 2002US-0404306P.  
XX PR 01-NOV-2002; 2002US-0423290P.  
XX PA (AGEN-) AGENSYS INC.

PI Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX Claim 1; Page 205; 443pp; English.  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
PS Sequence 9 AA;  
XX  
Query Match 93.3%; Score 28; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 1 PPPSY 5  
  
RESULT 7  
ADK88044  
ID ADK88044 standard; peptide; 9 AA.  
XX  
XX AC ADK88044;  
XX  
XX 20-MAY-2004 (first entry)  
XX Human 191P4D12(b) peptide fragment #4745.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
XX Homo sapiens.  
XX  
XX WO2004016799-A2.  
XX  
XX 26-FEB-2004.  
XX  
XX 23-APR-2003; 2003WO-US013013.  
XX  
XX 16-AUG-2002; 2002US-0404306P.  
XX  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX  
XX WPI; 2004-203808/19.  
XX  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX Claim 1; Page 199; 443pp; English.  
XX

CC The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
PS Sequence 9 AA;  
XX  
Query Match 93.3%; Score 28; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 5 PPPSY 9  
  
RESULT 8  
ADK84919  
ID ADK84919 standard; peptide; 9 AA.  
XX  
XX AC ADK84919;  
XX  
XX 20-MAY-2004 (first entry)  
XX Human 191P4D12(b) peptide fragment #1620.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
XX Homo sapiens.  
XX  
XX WO2004016799-A2.  
XX  
XX 26-FEB-2004.  
XX  
XX 23-APR-2003; 2003WO-US013013.  
XX  
XX 16-AUG-2002; 2002US-0404306P.  
XX  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX  
XX WPI; 2004-203808/19.  
XX  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
XX Claim 1; Page 160; 443pp; English.  
XX  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX



```
SQ Sequence 9 AA;
Query Match          93.3%; Score 28; DB 8; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 PPPXY 6
    |||||
Db 3 PPSY 7

RESULT 9
ADK84968
ID ADK84968 standard; peptide; 9 AA.
XX AC ADK84968;
XX DT 20-MAY-2004 (first entry)
XX DE Human 191P4D12(b) peptide fragment #1669.
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;
XX KW prostate cancer; bladder cancer; kidney cancer; colon cancer;
XX KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;
XX KW cervix cancer; immune response; human; chromosome 1q22-q23.2;
XX KW Suppression Subtractive Hybridisation; SSH.
XX OS Homo sapiens.
XX PN WO2004016799-A2.
XX PD 26-FEB-2004.
XX PF 23-APR-2003; 2003WO-US013013.
XX PR 16-AUG-2002; 2002US-0404306P.
XX PR 01-NOV-2002; 2002US-0423290P.
XX FA (AGEN-) AGENSYS INC.
XX PI Raitano AB, Challita-Bid PM, Jakobovits A, Faris M, Ge W;
XX WPI; 2004-203808/19.
XX PT New composition comprising 191P4D12(b) proteins and polynucleotides,
XX PT useful in diagnosing, preventing and treating cancer, e.g. prostate,
XX PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a
XX PT humoral or cellular immune response.
XX PS Claim 1; Page 160; 443pp; English.
XX CC The present invention relates to novel compositions comprising peptides
XX CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found
XX CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly
XX CC expressed in a number of cancers and so the compositions of the invention
XX CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of
XX CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,
XX CC uterus or cervix and in eliciting a humoral or cellular immune response.
XX CC To isolate genes that are overexpressed in prostate cancer, the
XX CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA
XX CC derived from prostate cancer tissues.

SQ Sequence 9 AA;
Query Match          93.3%; Score 28; DB 8; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 PPPXY 6
    |||||
Db 1 PPSY 5

RESULT 11
ADK84460
ID ADK84460 standard; peptide; 9 AA.
XX AC ADK84460;
XX DT 20-MAY-2004 (first entry)
XX DE Human 191P4D12(b) peptide fragment #1161.
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;

RESULT 10
ADK88384
ID ADK88384 standard; peptide; 9 AA.
XX AC ADK88384;
XX DT 20-MAY-2004 (first entry)
XX DE Human 191P4D12(b) peptide fragment #5085.
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;
XX KW prostate cancer; bladder cancer; kidney cancer; colon cancer;
XX KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;
XX KW cervix cancer; immune response; human; chromosome 1q22-q23.2;
XX KW Suppression Subtractive Hybridisation; SSH.
XX OS Homo sapiens.
XX PN WO2004016799-A2.
XX PD 26-FEB-2004.
XX PF 23-APR-2003; 2003WO-US013013.
XX PR 16-AUG-2002; 2002US-0404306P.
XX PR 01-NOV-2002; 2002US-0423290P.
XX FA (AGEN-) AGENSYS INC.
XX PI Raitano AB, Challita-Bid PM, Jakobovits A, Faris M, Ge W;
XX WPI; 2004-203808/19.
XX PT New composition comprising 191P4D12(b) proteins and polynucleotides,
XX PT useful in diagnosing, preventing and treating cancer, e.g. prostate,
XX PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a
XX PT humoral or cellular immune response.
XX PS Claim 1; Page 203; 443pp; English.
XX CC The present invention relates to novel compositions comprising peptides
XX CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found
XX CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly
XX CC expressed in a number of cancers and so the compositions of the invention
XX CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of
XX CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,
XX CC uterus or cervix and in eliciting a humoral or cellular immune response.
XX CC To isolate genes that are overexpressed in prostate cancer, the
XX CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA
XX CC derived from prostate cancer tissues.

SQ Sequence 9 AA;
Query Match          93.3%; Score 28; DB 8; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 PPPXY 6
    |||||
Db 3 PPSY 7

RESULT 11
ADK84460
ID ADK84460 standard; peptide; 9 AA.
XX AC ADK84460;
XX DT 20-MAY-2004 (first entry)
XX DE Human 191P4D12(b) peptide fragment #1161.
XX KW Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;
```

KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
OS Homo sapiens.  
XX  
PN WO2004016799-A2.  
XX  
XX 26-FEB-2004.  
XX  
XX 23-APR-2003; 2003WO-US013013.  
PF  
XX  
XX 16-AUG-2002; 2002US-0404306P.  
PR  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
PA  
XX Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
XX Claim 1; Page 155; 443pp; English.  
XX  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 93.3%; Score 28; DB 8; Length 9;  
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PPPXY 6  
Db 1 PPSY 5  
XX  
XX RESULT 12  
ADK86562  
ID ADK86562 standard; peptide; 9 AA.  
XX  
XX AC ADK86562;  
XX  
XX 20-MAY-2004 (first entry)  
XX  
XX Human 191P4D12(b) peptide fragment #3263.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
XX Homo sapiens.  
XX  
XX WO2004016799-A2.  
XX  
XX 26-FEB-2004.  
XX

PF prostate cancer; bladder cancer; kidney cancer; colon cancer;  
XX lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
PR cervix cancer; immune response; human; chromosome 1q22-q23.2;  
XX Suppression Subtractive Hybridisation; SSH.  
XX  
XX (AGEN-) AGENSYS INC.  
PA  
XX Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
XX Claim 1; Page 179; 443pp; English.  
XX  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 93.3%; Score 28; DB 8; Length 9;  
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PPPXY 6  
Db 3 PPSY 7  
XX  
XX RESULT 13  
ADK87588  
ID ADK87588 standard; peptide; 9 AA.  
XX  
XX AC ADK87588;  
XX  
XX 20-MAY-2004 (first entry)  
XX  
XX Human 191P4D12(b) peptide fragment #4289.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
XX Homo sapiens.  
XX  
XX WO2004016799-A2.  
XX  
XX 26-FEB-2004.  
XX  
XX 23-APR-2003; 2003WO-US013013.  
PF  
XX  
XX 16-AUG-2002; 2002US-0404306P.  
PR  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
PA  
XX Raitano AB, Challita-Eid PM, Jakobovits A, Faris M, Ge W;  
XX WPI; 2004-203808/19.  
XX

PT New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
PS Claim 1; Page 193; 443pp; English.  
XX  
CC The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
SQ Sequence 9 AA;  
Query Match 93.3%; Score 28; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PPPXY 6  
Db |||||  
2 PPSY 6  
RESULT 14  
ADK83902  
ID ADK83902 standard; peptide; 9 AA.  
XX  
AC ADK83902;  
XX  
XX 20-MAY-2004 (first entry)  
DT  
DE Human 191P4D12(b) peptide fragment #603.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
OS Homo sapiens.  
XX  
XX WO2004016799-A2.  
PN  
XX 26-FEB-2004.  
PD  
XX 23-APR-2003; 2003WO-US013013.  
PF  
XX 16-AUG-2002; 2002US-0404306P.  
PR  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Raitano AB, Challita-Bid PM, Jakobovits A, Faris M, Ge W;  
PI  
XX WPI; 2004-203808/19.  
DR  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
PS Claim 1; Page 148; 443pp; English.  
XX  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention

CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
SQ Sequence 9 AA;  
Query Match 93.3%; Score 28; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PPPXY 6  
Db |||||  
2 PPSY 6  
RESULT 15  
ADK88578  
ID ADK88578 standard; peptide; 9 AA.  
XX  
AC ADK88578;  
XX  
XX 20-MAY-2004 (first entry)  
DT  
DE Human 191P4D12(b) peptide fragment #5279.  
XX  
XX Cytostatic; Immunostimulant; Vaccine; Gene Therapy; 191P4D12(b); cancer;  
KW prostate cancer; bladder cancer; kidney cancer; colon cancer;  
KW lung cancer; pancreas cancer; ovary cancer; breast cancer; uterus cancer;  
KW cervix cancer; immune response; human; chromosome 1q22-q23.2;  
KW Suppression Subtractive Hybridisation; SSH.  
XX  
XX Homo sapiens.  
XX  
XX WO2004016799-A2.  
PN  
XX 26-FEB-2004.  
PD  
XX 23-APR-2003; 2003WO-US013013.  
PF  
XX 16-AUG-2002; 2002US-0404306P.  
PR  
XX 01-NOV-2002; 2002US-0423290P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Raitano AB, Challita-Bid PM, Jakobovits A, Faris M, Ge W;  
PI  
XX WPI; 2004-203808/19.  
DR  
XX New composition comprising 191P4D12(b) proteins and polynucleotides,  
PT useful in diagnosing, preventing and treating cancer, e.g. prostate,  
PT bladder, kidney, colon, lung, pancreas or ovary and in eliciting a  
PT humoral or cellular immune response.  
XX  
PS Claim 1; Page 206; 443pp; English.  
XX  
XX The present invention relates to novel compositions comprising peptides  
CC (ADK83300-ADK90584) from 191P4D12(b) and related proteins. It was found  
CC that 191P4D12(b), which maps to chromosome 1q22-q23.2, is aberrantly  
CC expressed in a number of cancers and so the compositions of the invention  
CC are useful in diagnosing, preventing and treating cancer, e.g. cancer of  
CC the prostate, bladder, kidney, colon, lung, pancreas, ovary, breast,  
CC uterus or cervix and in eliciting a humoral or cellular immune response.  
CC To isolate genes that are overexpressed in prostate cancer, the  
CC Suppression Subtractive Hybridisation (SSH) method was used, using cDNA  
CC derived from prostate cancer tissues.  
XX  
SQ Sequence 9 AA;  
Query Match 93.3%; Score 28; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6

Db 2 PPPSY 6

Search completed: October 13, 2005, 15:00:34  
Job time : 15.5165 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 2.85714 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-15  
Perfect score: 30  
Sequence: 1 XPPXPY 6

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	93.3	10	3	US-08-476-509B-35
2	28	93.3	14	3	US-08-630-916A-54
3	28	93.3	16	2	US-08-844-312-12
4	28	93.3	16	2	US-08-844-312-15
5	28	93.3	16	2	US-08-844-312-16
6	28	93.3	17	1	US-08-477-509B-19
7	28	93.3	17	1	US-08-482-085B-19
8	28	93.3	17	3	US-09-444-791A-19
9	28	93.3	21	3	US-08-867-611-55
10	28	93.3	21	4	US-09-690-359-55
11	28	93.3	32	3	US-09-314-268-156
12	28	93.3	52	4	US-09-270-767-35283
13	28	93.3	52	4	US-09-270-767-50500
14	28	93.3	61	4	US-09-513-999C-6691
15	28	93.3	63	3	US-09-154-083-22
16	28	93.3	64	4	US-09-328-352-6518
17	28	93.3	69	4	US-09-621-976-6022
18	28	93.3	74	4	US-09-270-767-33978
19	28	93.3	74	4	US-09-270-767-49195
20	28	93.3	75	4	US-09-270-767-58296
21	28	93.3	77	4	US-09-513-999C-6393
22	28	93.3	78	4	US-09-248-796A-25578
23	28	93.3	91	4	US-09-578-063-32
24	28	93.3	101	4	US-08-469-260A-48
25	28	93.3	101	4	US-08-488-446-48
26	28	93.3	101	4	US-08-467-344A-48
27	28	93.3	101	4	US-08-424-550B-48

28	28	93.3	106	4	US-09-640-211A-1103	Sequence 1103, Ap
29	28	93.3	115	4	US-09-248-796A-27933	Sequence 27933, A
30	28	93.3	119	4	US-09-270-767-40283	Sequence 40283, A
31	28	93.3	119	4	US-09-270-767-55499	Sequence 55499, A
32	28	93.3	128	4	US-09-902-540-10130	Sequence 10130, A
33	28	93.3	148	1	US-08-207-904-15	Sequence 15, Appl
34	28	93.3	149	4	US-03-489-039A-13180	Sequence 13180, A
35	28	93.3	150	4	US-09-270-767-59054	Sequence 59054, A
36	28	93.3	150	4	US-09-578-063-29	Sequence 29, Appl
37	28	93.3	150	4	US-09-578-063-44	Sequence 44, Appl
38	28	93.3	162	4	US-09-640-211A-2262	Sequence 2262, Ap
39	28	93.3	169	4	US-09-270-767-37020	Sequence 37020, A
40	28	93.3	169	4	US-03-270-767-52237	Sequence 52237, A
41	28	93.3	172	4	US-09-578-063-27	Sequence 27, Appl
42	28	93.3	172	4	US-09-578-063-74	Sequence 74, Appl
43	28	93.3	175	4	US-09-248-796A-16887	Sequence 16887, A
44	28	93.3	180	4	US-09-149-476-401	Sequence 401, App
45	28	93.3	187	3	US-09-247-155-102	Sequence 102, App

ALIGNMENTS

RESULT 1

US-08-476-509B-35  
; Sequence 35, Application US/08476509B  
; Patent No. 6034212  
; GENERAL INFORMATION:  
; APPLICANT: SUDOL, MARIUS  
; APPLICANT: PEER, BORK  
; APPLICANT: HENRY, CHEN  
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,509B  
; FILING DATE: 01-DEC-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEetical: NO  
US-08-476-509B-35  
Query Match 93.3%; Score 28; DB 3; Length 10;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 PPPXY 6
      |||||
Db      3 PPPAY 7

RESULT 2
US-08-630-916A-54
; Sequence 54, Application US/08630916A
; Patent No. 601137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630.916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-916A-54

Query Match      93.3%; Score 28; DB 3; Length 14;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
      |||||
Db      4 PPPAY 8

RESULT 3
US-08-844-312-12
; Sequence 12, Application US/08844312
; Patent No. 5948639
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639el TGF-(SYMBOL 98 \f "Symbol") Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

Query Match      93.3%; Score 28; DB 3; Length 14;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
      |||||
Db      4 PPPAY 8

RESULT 4
US-08-844-312-15
; Sequence 15, Application US/08844312
; Patent No. 5948639
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639el TGF-(SYMBOL 98 \f "Symbol") Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,312
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MN1-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-312-15

Query Match          93.3%; Score 28; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
Db 8 PPPAY 12

RESULT 5
US-08-844-312-16
; Sequence 16, Application US/08844312
; Patent No. 5948639
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639el TGF-{SYMBOL 98 \f "Symbol"} Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,312
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MMI-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-312-16

Query Match          93.3%; Score 28; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
Db 8 PPPAY 12

RESULT 6
US-08-477-509B-19
; Sequence 19, Application US/08477509B
; Patent No. 5770697
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; APPLICANT: Dorman, Mary A
; TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive
; MOLECULE TYPE: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,509B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-509B-19

Query Match          93.3%; Score 28; DB 1; Length 17;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
Db 11 PPPTY 15

RESULT 7
US-08-482-085B-19
; Sequence 19, Application US/08482085B
; Patent No. 6018030
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Richardson, Charles
; APPLICANT: Chambers, James
; APPLICANT: Causey, Stuart
; APPLICANT: Pollock, Thomas J.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W.
; TITLE OF INVENTION: No. 6018030el Peptides Comprising Repetitive
; UNITS OF AMINO ACIDS AND DNA SEQUENCES ENCODING THE SAME
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,085B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/927,258

FILING DATE: 04-NOV-1986

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/114,618

FILING DATE: 29-OCT-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/053,049

FILING DATE: 22-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,155

FILING DATE: 29-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-482-085B-19

Query Match 93.3%; Score 28; DB 3; Length 17;

Best Local Similarity 80.0%; Pred. No. 78;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PXPY 6

Db 11 PPTY 15

#### RESULT 8

US-09-444-791A-19

Sequence 19, Application US/09444791A

Patent No. 6355776

GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.

Richardson, Charles

Chambers, James

Casey, Stuart

Pollock, Thomas J.

Cappello, Joseph

Crisman, John W.

TITLE OF INVENTION: No. 6355776el Peptides Comprising Repetitive

Units of Amino Acids and DNA Sequences Encoding the Same

NUMBER OF SEQUENCES: 117

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/444,791A

FILING DATE: 22-No. 6355776-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/482,085

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/175,155

FILING DATE: 29-DEC-1993

APPLICATION NUMBER: US 08/053,049

FILING DATE: 22-APR-1993

APPLICATION NUMBER: US 07/114,618

FILING DATE: 29-OCT-1987

APPLICATION NUMBER: US 06/927,258

FILING DATE: 04-NOV-1986

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-444-791A-19

Query Match 93.3%; Score 28; DB 3; Length 17;

Best Local Similarity 80.0%; Pred. No. 78;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PXPY 6

Db 11 PPTY 15

#### RESULT 9

US-08-867-611-55

Sequence 55, Application US/08867611

Patent No. 6172189

GENERAL INFORMATION:

APPLICANT: DEVARE, SUSHIL G

APPLICANT: DESAI, SURESH M

APPLICANT: CASEY, JAMES M

APPLICANT: DAILEY, STEPHEN H

APPLICANT: DAWSON, GEORGE J

APPLICANT: GUTIERREZ, ROBIN A

APPLICANT: LESNIEWSKI, RICHARD R

APPLICANT: STEWART, JAMES L

APPLICANT: RUPPRECHT, KEVIN R

TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES

STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

CITY: ABBOTT PARK

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/667,611  
FILING DATE: 02-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,757  
FILING DATE:  
APPLICATION NUMBER: US/08/179,896  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/572,822  
FILING DATE: 24-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/614,069  
FILING DATE: 07-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/748,561  
FILING DATE: 21-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/748,565  
FILING DATE: 21-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/748,566  
FILING DATE: 21-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: FOREMSKI, PRISCILLA E  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 4834.US.P6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-9556  
TELEFAX: 708-937-9556  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-867-611-55

Query Match 93.3%; Score 28; DB 3; Length 21;  
Best Local Similarity 80.0%; Pred. No. 93;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
Db 1 PPSY 5

RESULT 10  
US-09-690-359-55  
Sequence 55, Application US/09690359  
Patent No. 6593083  
GENERAL INFORMATION:  
APPLICANT: DEVARE, SUSHIL G  
DESAI, SURESH M  
CASEY, JAMES M  
DALLEY, STEPHEN H  
DAWSON, GEORGE J  
GUTIERREZ, ROBIN A  
LESNIEWSKI, RICHARD R  
STEWART, JAMES L  
RUPPRECHT, KEVIN R  
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES  
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2  
CITY: ABBOTT PARK

STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/690,359  
FILING DATE: 17-Oct-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/867,611  
FILING DATE: 02-JUN-1997  
APPLICATION NUMBER: US/08/646,757  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/08/179,896  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/572,822  
FILING DATE: 24-AUG-1990  
APPLICATION NUMBER: US 07/614,069  
FILING DATE: 07-NOV-1990  
APPLICATION NUMBER: US 07/748,561  
FILING DATE: 21-AUG-1991  
APPLICATION NUMBER: US 07/748,565  
FILING DATE: 21-AUG-1991  
APPLICATION NUMBER: US 07/748,566  
FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:  
NAME: FOREMSKI, PRISCILLA E  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 4834.US.P6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 55:  
US-09-690-359-55

Query Match 93.3%; Score 28; DB 4; Length 21;  
Best Local Similarity 80.0%; Pred. No. 93;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
Db 1 PPSY 5

RESULT 11  
US-09-314-268-156  
Sequence 156, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doorbar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 156  
LENGTH: 39  
TYPE: PRT

; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: No. 6346377-Human Virus  
US-09-314-268-156

Query Match 93.3%; Score 28; DB 3; Length 39;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 10 PPPAY 14

RESULT 12  
US-09-270-767-35283  
; Sequence 35283, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35283  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-35283

Query Match 93.3%; Score 28; DB 4; Length 52;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 28 PPPTY 32

RESULT 13  
US-09-270-767-50500  
; Sequence 50500, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 50500  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-50500

Query Match 93.3%; Score 28; DB 4; Length 52;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 28 PPPTY 32

RESULT 14  
US-09-513-999C-6691  
; Sequence 6691, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6691  
; LENGTH: 61  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6691

Query Match 93.3%; Score 28; DB 4; Length 61;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 50 PPPAY 54

RESULT 15  
US-09-154-083-22  
; Sequence 22, Application US/09154083  
; Patent No. 6150513  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Kai  
; TITLE OF INVENTION: Polyketide Synthase Enzymes and Recombinant DNA  
; TITLE OF INVENTION: Constructs Therefor  
; FILE REFERENCE: 30062-20022.00 Polyketide Synthase Enz  
; CURRENT APPLICATION NUMBER: US/09/154.083  
; CURRENT FILING DATE: 1998-09-16  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Streptomyces hygroscopicus  
US-09-154-083-22

Query Match 93.3%; Score 28; DB 3; Length 63;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 PPPXY 6  
Db 51 PPPTY 55

Search completed: October 13, 2005, 14:03:06  
Job time : 3.98214 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 12.7912 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-15

Perfect score: 30

Sequence: 1 XPPXY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	93.3	5	14 US-10-226-007-5	Sequence 5, Appli
2	28	93.3	6	14 US-10-307-956-16	Sequence 16, Appl
3	28	93.3	8	14 US-10-226-007-25	Sequence 25, Appl
4	28	93.3	12	14 US-10-185-050-173	Sequence 173, App
5	28	93.3	13	14 US-10-185-050-168	Sequence 168, App
6	28	93.3	13	15 US-10-273-476-27	Sequence 27, Appl
7	28	93.3	14	14 US-10-185-050-54	Sequence 54, Appl
8	28	93.3	14	14 US-10-185-050-140	Sequence 140, App
9	28	93.3	14	14 US-10-307-956-17	Sequence 17, Appl
10	28	93.3	14	14 US-10-307-956-22	Sequence 22, Appl
11	28	93.3	14	14 US-10-307-956-23	Sequence 23, Appl

12	28	93.3	16	14 US-10-185-050-225	Sequence 225, App
13	28	93.3	16	14 US-10-185-050-229	Sequence 229, App
14	28	93.3	17	14 US-10-185-050-170	Sequence 170, App
15	28	93.3	17	14 US-10-185-050-171	Sequence 171, App
16	28	93.3	17	14 US-10-096-986-19	Sequence 19, Appl
17	28	93.3	20	18 US-10-862-195-196	Sequence 196, App
18	28	93.3	31	18 US-10-862-195-911	Sequence 911, App
19	28	93.3	39	14 US-10-008-5244-156	Sequence 156, App
20	28	93.3	39	15 US-10-350-719-156	Sequence 156, App
21	28	93.3	45	15 US-10-424-599-224011	Sequence 224011,
22	28	93.3	46	15 US-10-424-599-214590	Sequence 214590,
23	28	93.3	47	15 US-10-424-599-262825	Sequence 262825,
24	28	93.3	47	16 US-10-425-115-296064	Sequence 296064,
25	28	93.3	48	16 US-10-425-115-288558	Sequence 288558,
26	28	93.3	50	14 US-10-029-386-27771	Sequence 27771, A
27	28	93.3	51	15 US-10-437-963-107136	Sequence 107136,
28	28	93.3	51	15 US-10-424-599-170725	Sequence 170725,
29	28	93.3	51	16 US-10-425-115-318513	Sequence 318513,
30	28	93.3	52	16 US-10-425-115-316082	Sequence 316082,
31	28	93.3	57	15 US-10-242-355-400	Sequence 400, App
32	28	93.3	57	15 US-10-424-599-221354	Sequence 221354,
33	28	93.3	57	16 US-10-767-701-62862	Sequence 62862, A
34	28	93.3	58	16 US-10-437-963-165404	Sequence 165404,
35	28	93.3	59	15 US-10-424-599-254816	Sequence 254816,
36	28	93.3	59	15 US-10-424-599-258502	Sequence 258502,
37	28	93.3	59	16 US-10-425-115-222092	Sequence 222092,
38	28	93.3	62	20 US-11-097-143-29697	Sequence 29697, A
39	28	93.3	64	15 US-10-424-599-175735	Sequence 175735,
40	28	93.3	65	15 US-10-424-599-252903	Sequence 252903,
41	28	93.3	68	15 US-10-424-599-179260	Sequence 179260,
42	28	93.3	69	11 US-09-864-408A-6192	Sequence 6192, Ap
43	28	93.3	69	14 US-10-295-809-9	Sequence 9, Appl
44	28	93.3	69	16 US-10-425-115-268179	Sequence 268179,
45	28	93.3	70	16 US-10-437-963-168512	Sequence 168512,

#### ALIGNMENTS

##### RESULT 1

US-10-226-007-5  
; Sequence 5, Application US/10226007  
; Publication No. US20030105277A1  
; GENERAL INFORMATION:  
; APPLICANT: Myriad Genetics, Inc.  
; APPLICANT: Morham, Scott  
; APPLICANT: Zavitz, Kenton  
; APPLICANT: Hobden, Adrian  
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection  
; FILE REFERENCE: 5005.01  
; CURRENT APPLICATION NUMBER: US/10/226,007  
; CURRENT FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: US 60/313,883  
; PRIOR FILING DATE: 2001-08-21  
; NUMBER OF SEQ ID NOS: 1673  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-226-007-5

Query Match 93.3%; Score 28; DB 14; Length 5;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6

DB 1 PPPAY 5

##### RESULT 2

US-10-307-956-16

```
; Sequence 16, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-16

Query Match          93.3%; Score 28; DB 14; Length 6;
Best Local Similarity 80.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      2 PPPAY 6

RESULT 3
US-10-226-007-25
; Sequence 25, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-226-007-25

Query Match          93.3%; Score 28; DB 14; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      1 PPPAY 5

RESULT 4
US-10-185-050-173
; Sequence 173, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-10-185-050-173
```

```
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-10-185-050-173

Query Match          93.3%; Score 28; DB 14; Length 12;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      8 PPPAY 12

RESULT 5
US-10-185-050-168
; Sequence 168, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-10-185-050-173
```

```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-10-185-050-168

Query Match          93.3%; Score 28; DB 14; Length 13;
Best Local Similarity 80.0%; Pred No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PPPXY 6
Db      4 PPPAY 8

RESULT 6
US-10-273-476-27
; Sequence 27, Application US/10273476
; Publication No. US20040023234A1
; GENERAL INFORMATION:
; APPLICANT: Hildebrandt, Friedhelm
; APPLICANT: Jentsch, Thomas J.
; TITLE OF INVENTION: BNSD Nucleic Acids and Proteins
; FILE REFERENCE: UM-07447
; CURRENT APPLICATION NUMBER: US/10/273,476
; CURRENT FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-476-27

Query Match          93.3%; Score 28; DB 15; Length 13;
Best Local Similarity 80.0%; Pred No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PPPXY 6
Db      5 PPPAY 9

RESULT 7
US-10-185-050-54
; Sequence 54, Application US/10185050
; Publication No. US2003007577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas

```

```

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/185,050
FILING DATE: 28-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/826,516
FILING DATE: 03-Apr-1997
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-208-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-185-050-54

Query Match          93.3%;   Score 28;   DB 14;   Length 14;
Best Local Similarity 80.0%;   Pred No. 4.9e+02;
Matches 4;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY      2 PPPXY 6
      |||||
Db      4 PPPAY 8

RESULT 8
US-10-185-050-140
: Sequence 140, Application US/10185050
: Publication No. US20030077577A1
: GENERAL INFORMATION:
: APPLICANT: Pirozzi, Gregorio
: Kay, Brian K.
: Fowlkes, Dana M.
: TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
:
: NUMBER OF SEQUENCES: 233
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/185,050
: FILING DATE: 28-Jun-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/826,516
: FILING DATE: 03-Apr-1997

```

```

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/185,050
FILING DATE: 28-Jun-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/826,516
FILING DATE: 03-Apr-1997
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-208-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-185-050-54

Query Match          93.3%;   Score 28;   DB 14;   Length 14;
Best Local Similarity 80.0%;   Pred No. 4.9e+02;
Matches 4;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY      2 PPPXY 6
      |||||
Db      4 PPPAY 8

RESULT 8
US-10-185-050-140
: Sequence 140, Application US/10185050
: Publication No. US20030077577A1
: GENERAL INFORMATION:
: APPLICANT: Pirozzi, Gregorio
: Kay, Brian K.
: Fowlkes, Dana M.
: TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
:
: NUMBER OF SEQUENCES: 233
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/185,050
: FILING DATE: 28-Jun-2002
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/826,516
: FILING DATE: 03-Apr-1997

```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 140:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 140:
US-10-185-050-140

Query Match          93.3%; Score 28; DB 14; Length 14;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      5 PPPAY 9

RESULT 9
US-10-307-956-17
; Sequence 17, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (10)...(10)
; OTHER INFORMATION: Xaa = Leucine or Methionine
US-10-307-956-17

Query Match          93.3%; Score 28; DB 14; Length 14;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      5 PPPAY 9

RESULT 10
US-10-307-956-22
; Sequence 22, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
```

```
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-22

Query Match          93.3%; Score 28; DB 14; Length 14;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      5 PPPAY 9

RESULT 11
US-10-307-956-23
; Sequence 23, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-23

Query Match          93.3%; Score 28; DB 14; Length 14;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PPPXY 6
Db      5 PPPAY 9

RESULT 12
US-10-185-050-225
; Sequence 225, Application US/10185050
; Publication No. US2003007577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/185,050  
FILING DATE: 28-Jun-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,516  
FILING DATE: 03-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-208-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
US-10-185-050-225

Query Match 93.3%; Score 28; DB 14; Length 16;  
Best Local Similarity 80.0%; Pred. No. 5.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
Db 6 PPSY 10

RESULT 13  
US-10-185-050-229  
Sequence 229, Application US/10185050  
Publication No. US20030077577A1  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
Kay, Brian K.  
Fowlkes, Dana M.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
NUMBER OF SEQUENCES: 233  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/185,050  
FILING DATE: 28-Jun-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,516  
FILING DATE: 03-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-208-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 229:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 229:  
US-10-185-050-229

Query Match 93.3%; Score 28; DB 14; Length 16;  
Best Local Similarity 80.0%; Pred. No. 5.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
Db 6 PPSY 10

RESULT 14  
US-10-185-050-170  
Sequence 170, Application US/10185050  
Publication No. US20030077577A1  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
Kay, Brian K.  
Fowlkes, Dana M.  
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
NUMBER OF SEQUENCES: 233  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/185,050  
FILING DATE: 28-Jun-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,516  
FILING DATE: 03-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-208-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 170:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 170:  
US-10-185-050-170

Query Match 93.3%; Score 28; DB 14; Length 17;  
Best Local Similarity 80.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPXY 6  
| | | |  
Db 8 PPAY 12

## RESULT 15

US-10-185-050-171  
; Sequence 171, Application US/10185050  
; Publication No. US2003007577A1  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; Kay, Brian K.  
; Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 233  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/185,050  
; FILING DATE: 28-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/826,516  
; FILING DATE: 03-Apr-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-208-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 171:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 171:  
US-10-185-050-171

Query Match 93.3%; Score 28; DB 14; Length 17;  
Best Local Similarity 80.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPXY 6  
| | | |  
Db 8 PPAY 12

Search completed: October 13, 2005, 14:50:02  
Job time : 13.7912 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:36:01 ; Search time 2.50549 Seconds  
(without alignments)  
230.414 Million cell updates/sec

Title: US-09-385-918-15  
Perfect score: 30  
Sequence: 1 XPPXY 6  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	93.3	62	S74359	hypothetical prote
2	28	93.3	66	T31844	hypothetical prote
3	28	93.3	70	T31845	hypothetical prote
4	28	93.3	82	T33088	hypothetical prote
5	28	93.3	82	T41925	hypothetical prote
6	28	93.3	85	T05386	hypothetical prote
7	28	93.3	88	T44170	hypothetical prote
8	28	93.3	88	T43984	hypothetical prote
9	28	93.3	93	T09318	BoLF1 protein - hu
10	28	93.3	99	T33486	hypothetical prote
11	28	93.3	105	E64498	hypothetical prote
12	28	93.3	115	AD2303	hypothetical prote
13	28	93.3	118	W4WL82	Pex protein (impor
14	28	93.3	132	S14970	E4 protein - bovin
15	28	93.3	143	S42579	extensin class I (
16	28	93.3	143	T29840	QID3 protein - fun
17	28	93.3	148	D72652	hypothetical prote
18	28	93.3	149	T41481	very hypothetical
19	28	93.3	150	T39811	hypothetical prote
20	28	93.3	153	P96575	hypothetical prote
21	28	93.3	155	D82755	conserved hypothet
22	28	93.3	163	A29356	hydroxyproline-ric
23	28	93.3	164	H87551	conserved hypothet
24	28	93.3	166	T29641	hypothetical prote
25	28	93.3	167	AC3314	hypothetical membr
26	28	93.3	170	A72638	hypothetical prote
27	28	93.3	171	T20598	hypothetical prote
28	28	93.3	171	T25696	hypothetical prote
29	28	93.3	172	T23261	hypothetical prote

30	28	93.3	172	2	T27505	hypothetical prote
31	28	93.3	177	2	T18677	hypothetical prote
32	28	93.3	178	2	T33440	hypothetical prote
33	28	93.3	191	2	E87601	OmpA family protei
34	28	93.3	198	2	D70509	hypothetical prote
35	28	93.3	199	1	S31938	membrane-bound cyt
36	28	93.3	200	2	T34104	hypothetical prote
37	28	93.3	203	2	S51921	signal peptidase I
38	28	93.3	210	2	T04233	pathogenesis-relat
39	28	93.3	210	2	G97235	hypothetical prote
40	28	93.3	212	2	F72579	hypothetical prote
41	28	93.3	219	2	S35643	BTEB2 protein - hu
42	28	93.3	222	2	S33204	hypothetical prote
43	28	93.3	223	2	T29976	hypothetical prote
44	28	93.3	223	2	T16654	hypothetical prote
45	28	93.3	224	2	T37825	serine /proline ri

ALIGNMENTS

RESULT 1  
S74359  
hypothetical protein ssr0109 - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S74359  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
S.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74359  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-62 <KAN>  
A:Cross-references: UNIPROT:P72574; EMBL:D64001; GB:AB001339; NID:g1001102; PIDN:BAA102  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 93.3%; Score 28; DB 2; Length 62;  
Best Local Similarity 80.0%; Pred. No. 1.le+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PPPXY 6  
Db 20 PPPTY 24

RESULT 2  
T31844  
hypothetical protein C02E7.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T31844  
R:Fulton, B.; Wohldmann, P.  
submitted to the EMBL Data Library, June 1998  
A:Description: The sequence of C. elegans cosmid C02E7.  
A:Reference number: Z21093  
A:Accession: T31844  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-66 <FUL>  
A:Cross-references: UNIPROT:O16427; EMBL:AF016446; PIDN:AAC24166.1; GSPDB:GN00023; CESP  
A:Experimental source: strain Bristol N2; clone C02E7  
C:Genetics:  
A:Gene: CESP:C02E7.7  
A:Map position: 5  
A:Introns: 29/2

Query Match 93.3%; Score 28; DB 2; Length 66;  
Best Local Similarity 80.0%; Pred. No. 1.le+02;

```
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
   |||||
Db 50 PPPAY 54

RESULT 3
T31845
hypotheical protein C02E7.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31845
R:Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, June 1998
A:Description: The sequence of C. elegans cosmid C02E7.
A:Reference number: Z21093
A:Accession: T31845
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-70 <FUL>
A:Cross-references: UNIPROT:O16428; EMBL:AF016446; PIDN:AAC24165.1; GSPDB:GN00023; CESP:
A:Experimental source: strain Bristol N2; clone C02E7
C:Genetics:
A:Gene: CESP:C02E7.6
A:Map position: 5
A:Introns: 29/2

Query Match 93.3%; Score 28; DB 2; Length 70;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
   |||||
Db 45 PPPAY 49

RESULT 4
T33088
hypotheical protein R12E2.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33088
R:Goala, D.; Scheer, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid R12E2.
A:Reference number: Z21281
A:Accession: T33088
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-82 <GOE>
A:Cross-references: UNIPROT:O61784; EMBL:AF067219; PIDN:AAC17028.1; GSPDB:GN00019; CESP:
A:Experimental source: strain Bristol N2; clone R12E2
C:Genetics:
A:Gene: CESP:R12E2.7
A:Map position: 1
A:Introns: 41/2

Query Match 93.3%; Score 28; DB 2; Length 82;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
   |||||
Db 65 PPPAY 69

RESULT 5
T41925
hypotheical protein U24 - human herpesvirus 7 (strain JI)
C:Species: human herpesvirus 7
A:Variety: strain JI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
```

```
C:Accession: T41925
R:Nicholas, J.
submitted to the EMBL Data Library, December 1995
A:Description: Determination and analysis of the complete nucleotide sequence of human
A:Reference number: Z22022
A:Accession: T41925
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-82 <NIC>
A:Cross-references: UNIPROT:Q69505; EMBL:U43400; PIDN:AAC54685.1
A:Experimental source: strain JI
C:Genetics:
A:Note: U24

Query Match 93.3%; Score 28; DB 2; Length 82;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
   |||||
Db 6 PPPSY 10

RESULT 6
T05386
hypotheical protein F16G20.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05386
R:Bevan, M.; Watson, M.D.; Gibbons, T.; Bartley, J.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, August 1998
A:Reference number: Z15413
A:Accession: T05386
A:Molecule type: DNA
A:Residues: 1-85 <BEV>
A:Cross-references: UNIPROT:O81744; EMBL:AL031326
A:Experimental source: cultivar Columbia; BAC clone F16G20
C:Genetics:
A:Map position: 4
A:Note: F16G20.180

Query Match 93.3%; Score 28; DB 2; Length 85;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6
   |||||
Db 79 PPPAY 83

RESULT 7
T44170
hypotheical protein U24 [imported] - human herpesvirus 6 (strain Z29)
C:Species: human herpesvirus 6
A:Variety: strain Z29
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Jun-2000
C:Accession: T44170
R:Dominguez, G.; Dambagh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
J. Virol. 73, 8040-8052, 1999
A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with human
A:Reference number: Z22734; MUID:99412318; PMID:10482553
A:Accession: T44170
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-88 <DOM>
A:Cross-references: EMBL:AF157706; PIDN:AAD49637.1
A:Experimental source: strain Z29; variant B
C:Genetics:
A:Note: U24

Query Match 93.3%; Score 28; DB 2; Length 88;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 2 PPPXY 6  
|||  
Db 7 PPSY 11

## RESULT 8

T43984  
Hypothetical protein U24 [imported] - human herpesvirus 6 (strain HST)  
C:Species: human herpesvirus 6  
A:Variety: strain HST  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T43984  
R:Riggawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kaw  
J. Virol. 73, 8053-8063, 1999  
A:Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and  
A:Reference number: Z22732; MUID:99412319; PMID:10482554  
A:Accession: T43984  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-88 <ISE>  
A:Cross-references: UNIPROT:Q9WT40; EMBL:AB021506; NID:g4995977; PIDN:BAA78245.1; PID:g4  
A:Experimental source: strain HST; pop. variant B  
C:Genetics:  
A:Note: U24

Query Match 93.3%; Score 28; DB 2; Length 88;  
Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
|||  
Db 7 PPSY 11

## RESULT 9

T09318  
EoLF1 protein - human herpesvirus 6 (strain UI102)  
C:Species: human herpesvirus 6  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09318  
R:Nicholas, J.; Martin, M.  
J. Virol. 68, 597-610, 1994  
A:Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of hu  
A:Reference number: Z16644; MUID:94118404; PMID:8289364  
A:Accession: T09318  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-93 <NIC>  
A:Cross-references: UNIPROT:Q69048; EMBL:L25528; NID:g451932; PIDN:AAA16731.1; PID:g4519  
C:Genetics:  
A:Gene: EoLF1

Query Match 93.3%; Score 28; DB 2; Length 93;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
|||  
Db 13 PPSY 17

## RESULT 10

T33486  
Hypothetical protein C08G5.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33486  
R:Klinka, W.; Bauer, C.; Morris, M.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid C08G5.  
A:Reference number: Z21356  
A:Accession: T33486

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-99 <KLI>  
A:Cross-references: UNIPROT:Q9TZL2; EMBL:AF098985; PIDN:AAC67418.1; GSPDB:GN00020; CESP  
A:Experimental source: strain Bristol N2; clone C08G5  
C:Genetics:  
A:Gene: CESP:C08G5.5  
A:Map position: 2  
A:Introns: 30/3; 75/3

Query Match 93.3%; Score 28; DB 2; Length 99;  
Best Local Similarity 80.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
|||  
Db 81 PPSY 85

## RESULT 11

E64498  
Hypothetical protein MJ1590 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: E64498  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.,  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasch  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: E64498  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-105 <BUL>  
A:Cross-references: UNIPROT:Q58985; GB:U67599; GB:L77117; NID:g2826435; PIDN:AAB99618.1,  
C:Genetics:  
A:Map position: FOR1562771-1563088

Query Match 93.3%; Score 28; DB 2; Length 105;  
Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
|||  
Db 27 PPSY 31

## RESULT 12

AD2303  
Pex protein [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AD2303  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, A.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD2303  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-115 <KUR>  
A:Cross-references: UNIPROT:Q8Y056; GB:BA0000019; PIDN:BA875678.1; PID:g17133113; GSPDB:  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: pex

Query Match 93.3%; Score 28; DB 2; Length 115;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 |||||  
 Db 13 PPPTY 17

## RESULT 13

W4WLB2

E4 protein - bovine papillomavirus type 2

C:Species: bovine papillomavirus type 2

C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004

C:Accession: F31169

R:Groff, D.E.; Mitra, R.; Lancaster, W.D.

submitted to GenBank, May 1988

A:Reference number: A94519

A:Accession: F31169

A:Molecule type: DNA

A:Residues: 1-118 &lt;GRO&gt;

A:Cross-references: UNIPROT:P11301; GB:M20219; GB:M19551; NID:g332996

C:Superfamily: bovine papillomavirus E4 protein

C:Keywords: early protein

Query Match 93.3%; Score 28; DB 1; Length 118;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 |||||  
 Db 51 PPAY 55

## RESULT 14

S14970

extensin class I (clone w17-1) - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S14970

R:Showalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.

Plant Mol. Biol. 16, 547-565, 1991

A:Title: Tomato extensin and extensin-like cDNAs: structure and expression in response to

A:Reference number: S14970; MUID:91329690; PMID:1714316

A:Accession: S14970

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-132 &lt;SHO&gt;

A:Cross-references: UNIPROT:Q01942; EMBL:X55681; NID:g19212; PIDN:CAA39212.1; PID:g13455

C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 93.3%; Score 28; DB 2; Length 132;  
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 |||||  
 Db 94 PPPTY 98

## RESULT 15

S42579

QID3 protein - fungus (Trichoderma harzianum)

C:Species: Trichoderma harzianum

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S42579

R:Lora, J.M.; de la Cruz, J.; Benitez, T.; Llobell, A.; Pintor-Toro, J.A.

Mol. Gen. Genet. 242, 461-466, 1994

A:Title: A putative catabolite-repressed cell wall protein from the mycoparasitic fungus

A:Reference number: S42579; MUID:94166756; PMID:8121402

A:Accession: S42579

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-143 &lt;LOR&gt;

A:Cross-references: UNIPROT:P52755; EMBL:X71913; NID:g460818; PIDN:CAA50728.1; PID:g4608

C:Superfamily: hydrophobin HFBI

Query Match 93.3%; Score 28; DB 2; Length 143;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 |||||  
 Db 25 PPPTY 29

Search completed: October 13, 2005, 15:11:18

Job time : 5.50549 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 11.4286 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-15

Perfect score: 30

Sequence: 1 XPPPY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_prot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	93.3	21	2 Q9M3S8	Q9M3S8 arabidopsis
2	28	93.3	25	2 Q8WM19	Q8WM19 macaca mula
3	28	93.3	37	2 Q96AU8	Q96AU8 homo sapien
4	28	93.3	38	2 Q8HY43	Q8HY43 bos taurus
5	28	93.3	42	2 Q8AV30	Q8AV30 brachydanio
6	28	93.3	42	2 Q04217	Q04217 bromheadia
7	28	93.3	62	2 P72574	P72574 synechocyst
8	28	93.3	64	2 Q943H5	Q943H5 oryza sativ
9	28	93.3	66	2 Q16427	Q16427 caenorhabdi
10	28	93.3	69	2 Q9VPF8	Q9VPF8 human herpe
11	28	93.3	70	2 Q16428	Q16428 caenorhabdi
12	28	93.3	71	2 Q04216	Q04216 bromheadia
13	28	93.3	73	2 Q65Q22	Q65Q22 oryza sativ
14	28	93.3	75	2 Q7XN09	Q7XN09 oryza sativ
15	28	93.3	75	2 Q88460	Q88460 mus musculu
16	28	93.3	76	1 SP2E MOUSE	Q70556 mus musculu
17	28	93.3	76	1 SP2I MOUSE	Q70560 mus musculu
18	28	93.3	78	2 Q9BL10	Q9BL10 caenorhabdi
19	28	93.3	82	2 Q61784	Q61784 caenorhabdi
20	28	93.3	82	2 Q01947	Q01947 lycopersico
21	28	93.3	82	2 Q69505	Q69505 human herpe
22	28	93.3	85	2 Q81744	Q81744 arabidopsis
23	28	93.3	87	2 Q69559	Q69559 human herpe
24	28	93.3	88	2 Q94ES2	Q94ES2 pisum sativ
25	28	93.3	88	2 Q9QU42	Q9QU42 human herpe
26	28	93.3	88	2 Q9WT40	Q9WT40 human herpe
27	28	93.3	92	2 Q98626	Q98626 pneumonia v
28	28	93.3	92	2 Q6PWL0	Q6PWL0 pneumonia v
29	28	93.3	93	2 Q7VTJ9	Q7VTJ9 caenorhabdi
30	28	93.3	93	2 Q69048	Q69048 human herpe
31	28	93.3	94	2 Q865B4	Q865B4 equus cabal

32	28	93.3	97	2 Q7S376	Q7S376 neurospora
33	28	93.3	98	2 Q6IL41	Q6IL41 drosophila
34	28	93.3	98	2 Q748N0	Q748N0 geobacter s
35	28	93.3	99	2 Q9TZL2	Q9TZL2 caenorhabdi
36	28	93.3	101	2 Q8CEH0	Q8CEH0 mus musculu
37	28	93.3	102	2 Q8S5G3	Q8S5G3 oryza sativ
38	28	93.3	102	2 Q7G6H0	Q7G6H0 oryza sativ
39	28	93.3	103	2 Q9BGZ7	Q9BGZ7 macaca fasc
40	28	93.3	105	1 YF90_METJA	Q58985 methanococc
41	28	93.3	107	2 Q6FM10	Q6FM10 candida gla
42	28	93.3	108	1 SP2H_MOUSE	Q70559 mus musculu
43	28	93.3	109	2 Q9ELB3	Q9ELB3 caenorhabdi
44	28	93.3	111	2 Q96KD1	Q96KD1 homo sapien
45	28	93.3	112	1 VEA_BP2	P11301 bovine papi

## ALIGNMENTS

RESULT 1					
Q9M3S8					
ID	Q9M3S8	PRELIMINARY;	PRT;	21 AA.	
AC	Q9M3S8;				
DT	01-OCT-2000 (Tremblrel. 15, Created)				
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)				
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)				
DE	Extensin-like protein (Fragment).				
GN	Name=DiDi 4T-2;				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.				
OX	NCBI_TaxID=3702;				
RN	[1]_SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Roots;				
RX	MEDLINE=21171025; PubMed=11277426;				
RA	Vercauteren I., Van Der Schueren B., Van Montagu M., Gheysen G.;				
RT	"Arabidopsis thaliana genes expressed in the early compatible				
RT	interaction with root-knot nematodes.";				
RL	Mol. Plant Microbe Interact. 14:288-299 (2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Roots;				
RA	Vercauteren I.J.R.;				
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AJ286344; CAB71008.1;				
FT	NON_TER				
SQ	SEQUENCE 21 AA; 2320 MW; 3D31549B456A4726 CRC64;				
Query Match 93.3%; Score 28; DB 2; Length 21;					
Best Local Similarity 80.0%; Pred. No. 2.5e+02;					
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	2 PPPXY 6				
Db	9 PPPSY 13				
RESULT 2					
Q8WM19					
ID	Q8WM19	PRELIMINARY;	PRT;	25 AA.	
AC	Q8WM19;				
DT	01-MAR-2002 (Tremblrel. 20, Created)				
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)				
DT	01-MAR-2002 (Tremblrel. 20, Last annotation update)				
DE	Vesicular glutamate transporter 1 (Fragment).				
OS	Macaca mulatta (Rhesus macaque).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;				
OC	Cercopitheciinae; Macaca.				
OX	NCBI_TaxID=9544;				
RN	[1]				

RP SEQUENCE FROM N.A.  
 RC TISSUE=Medial basal hypothalamus;  
 RA Brown A.E., Ojeda S.R.; to the EMBL/GenBank/DBDJ databases.  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBDJ databases.  
 DR EMBL; AF424830; AAL40244.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 25 AA; 2733 MW; 1730048E8674A5F3 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
 |||||  
 Db 2 PPSY 6

RESULT 3  
 Q96AU8 PRELIMINARY; PRT; 37 AA.  
 AC Q96AU8;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBDJ databases.  
 DR EMBL; BC016719; AAH16719.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 37 AA; 4001 MW; A1EAA3B5B5E6734C CRC64;

Query Match 93.3%; Score 28; DB 2; Length 37;  
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
 |||||  
 Db 12 PPPAY 16

RESULT 4  
 Q8HY43 PRELIMINARY; PRT; 38 AA.  
 AC Q8HY43;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Early growth response protein 1 (Fragment).  
 GN Name=egr1;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Prigmet C., Sarikaya H., Bruckmaier R.M., Pfaffl M.W.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBDJ databases.  
 DR EMBL; AJ535318; CAD59473.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 38  
 SQ SEQUENCE 38 AA; 4064 MW; CAE0F5FD241DC202 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 38;

Best Local Similarity 80.0%; Pred. No. 4.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 PPPXY 6  
 |||||  
 Db 2 PPPAY 6

RESULT 5  
 Q8AV30 PRELIMINARY; PRT; 42 AA.  
 AC Q8AV30;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Smad5 (Fragment).  
 GN Name=smad5;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2263387; PubMed=12376102; DOI=10.1016/S0012-1606(02)90805-5;  
 RA Kramer C., Mayr T., Nowak M., Bauer H., Schumacher J., Runke G.,  
 RA Wagner D.S., Schmid B., Imai Y., Talbot W.S., Mullins M.C.,  
 RA Hammerschmidt M.;  
 RT "Maternally supplied Smad5 is required for ventral specification in  
 RT zebrafish embryos prior to zygotic Bmp signaling.";  
 RL Dev. Biol. 250:263-279(2002).  
 DR EMBL; AY135144; AAN34935.1; -.  
 DR EMBL; AY135143; AAN34935.1; JOINED.  
 FT NON\_TER 1  
 FT NON\_TER 42  
 SQ SEQUENCE 42 AA; 4263 MW; B1E82C53C0A7A01E CRC64;

Query Match 93.3%; Score 28; DB 2; Length 42;  
 Best Local Similarity 80.0%; Pred. No. 5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PPPXY 6  
 |||||  
 Db 28 PPPAY 32

RESULT 6  
 O04217 PRELIMINARY; PRT; 48 AA.  
 AC O04217;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Extensin (Fragment).  
 OS Bromheadia finlaysoniana (Orchid).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;  
 OC Epidendroideae; higher Epidendroideae; Cymbidieae; Bromheadiinae;  
 OC Bromheadia.  
 OX NCBI\_TaxID=41205;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Floral;  
 RA Lim S.H.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBDJ databases.  
 DR EMBL; Y13142; CAA73602.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 48 AA; 5317 MW; C61161A476AEB6CF CRC64;

Query Match 93.3%; Score 28; DB 2; Length 48;  
 Best Local Similarity 80.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 Db 33 PPPTY 37

## RESULT 7

P72574 PRELIMINARY; PRT; 62 AA.  
 AC P72574;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ssr0109 protein.  
 GN OrderedLocuNames=ssr0109;  
 OS Synchocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]\_TaxID=1148;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC6803;  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirose M., Sugiura M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res 3:109-136(1996).  
 DR EMBL; D64001; BAA10277.1; --  
 DR PIR; S74359; S74359.  
 KW Complete proteome.  
 SQ SEQUENCE 62 AA; 6743 MW; F9ABB2434C0C9406 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 62;  
 Best Local Similarity 80.0%; Pred. No. 7.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 Db 20 PPPTY 24

## RESULT 8

Q943H5 PRELIMINARY; PRT; 64 AA.  
 AC Q943H5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE P0046E05.3 protein.  
 GN Name=P0046E05.3;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]\_TaxID=39947;  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12447438; DOI=10.1038/nature01184;  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayaishi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K. Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,

RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL; AP003237; BAB67898.1; --  
 DR Gramene; Q943H5; --  
 SQ SEQUENCE 64 AA; 7001 MW; 8E5A64A20BC656C9 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 64;  
 Best Local Similarity 80.0%; Pred. No. 7.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 Db 18 PPPTY 22

## RESULT 9

O16427 PRELIMINARY; PRT; 66 AA.  
 AC O16427;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein C02E7.7.  
 GN Name=C02E7.7; ORFNames=C02E7.7;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG WormBase Consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. the C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Fulton B., Wohldmann P.;  
 RT "The sequence of C. elegans cosmid C02E7.";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF016446; AAC24166.1; --  
 DR PIR; T31844; T31844.  
 DR WormBase; WBGene00015340; C02E7.7.  
 DR WormPep; C02E7.7; CE07844.  
 KW Hypothetical protein.  
 SQ SEQUENCE 66 AA; 6077 MW; E393B3D0B8E39915 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 66;  
 Best Local Similarity 80.0%; Pred. No. 7.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPPXY 6  
 Db 50 PPPTY 54

```

RESULT 10
Q9YPF8      PRELIMINARY;      PRT;      69 AA.
ID Q9YPF8;
AC Q9YPF8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Latency associated transcript
OS Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86155754; PubMed=28311380;
RA Wagner E.K., Devi-Rao G.B., Feldman L.T., Dobson A.T., Zhang Y.-K.,
RA Flanagan W.M., Stevens J.G.;
RT "Physical characterization of the herpes simplex virus latency-
RT associated transcript in neurons.";
RL J. Virol. 62:1194-1202(1988).
DR EMBL; M17921; AAA45796.1; -.
SQ SEQUENCE 69 AA; 7286 MW; 9E26E5B0750DB39A CRC64;

Query Match      93.3%; Score 28; DB 2; Length 69;
Best Local Similarity 80.0%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PPPXY 6
Db      |||||
        56 PPSY 60

RESULT 11
O16428      PRELIMINARY;      PRT;      70 AA.
ID O16428;
AC O16428;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein C02E7.6.
GN Name=C02E7.6; ORFNames=C02E7.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=98519116;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton B., Woldmann P.;
RT "The sequence of C. elegans cosmid C02E7.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016446; AAC24165.1; -.
DR PIR; T31845; T31845.
DR WormBase; WBGene00015339; C02E7.6.
DR WormPep; C02E7.6; CE07843.
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 6547 MW; 9FE8F12AA71C46FD CRC64;

Query Match      93.3%; Score 28; DB 2; Length 70;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PPPXY 6
Db      |||||
        45 PPAY 49

RESULT 12
O04216      PRELIMINARY;      PRT;      71 AA.
ID O04216;
AC O04216;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extensin (Fragment).
OS Bromheadia finlaysoniana (Orchid).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Epidendroideae; higher Epidendroideae; Cymbidieae; Bromheadiinae;
OC Bromheadia.
OX NCBI_TaxID=41205;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Floral;
RA Lim S.H.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13141; CAA73601.1; -.
FT NON TER 1
SQ SEQUENCE 71 AA; 7715 MW; 5FEF6A5483C9DD8C CRC64;

Query Match      93.3%; Score 28; DB 2; Length 71;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PPPXY 6
Db      |||||
        56 PPPT 60

RESULT 13
Q06QZ2      PRELIMINARY;      PRT;      73 AA.
ID Q06QZ2;
AC Q06QZ2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSUNBa0055113.11.
GN Name=OSUNBa0055113.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005648; BAD28928.1; -.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 7920 MW; F3EB1BAF0E1FA1CB CRC64;

```



Query Match 93.3%; Score 28; DB 2; Length 73;  
Best Local Similarity 80.0%; Pred. No. 8e+02; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 2 PPPXY 6  
|||  
DB 6 PPPTY 10

## RESULT 14

Q7XN09 PRELIMINARY; PRT; 75 AA.  
AC Q7XN09;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE OSJNBa0008M17.8 protein.  
GN OSJNBa0008M17.8;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=1247439; DOI=10.1038/nature01183;  
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang Y.,  
Wang Q., Zhang L., Lu Y., Mu Y., Zhang L.S., Yu Z., Fan D.,  
Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
Han B.;  
RT "Sequence and analysis of rice chromosome 4."  
RL Nature 420:316-320(2002).  
DR EMBL; AL662950; CAE04335.2; -.  
DR Gramene; Q7XN09; -.  
SQ SEQUENCE 75 AA; 7723 MW; A88ECB2376BB373D CRC64;

Query Match 93.3%; Score 28; DB 2; Length 75;  
Best Local Similarity 80.0%; Pred. No. 9e+02; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 2 PPPXY 6  
|||  
DB 21 PPPAY 25

## RESULT 15

O88460 PRELIMINARY; PRT; 75 AA.  
AC O88460;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE EGF-like growth factor receptor ErbB4 intracellular domain  
(Fragment).  
GN Name=ErbB4;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CD-1; TISSUE=Uterus;  
RA Lim H., Das S.K., Dey S.K.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF059177; AAC28334.1; -.  
DR MGD; MGI:104771; ErbB4.  
DR GO; GO:0045165; P:cell fate commitment; IDA.

DR GO; GO:0007507; P:heart development; IMP.  
DR GO; GO:0007399; P:neurogenesis; IMP.  
KW Receptor.  
FT NON\_TER 1  
FT NON\_TER 75  
SQ SEQUENCE 75 AA; 8371 MW; 718C044E67673A70 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 75;  
Best Local Similarity 80.0%; Pred. No. 9e+02; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 2 PPPXY 6  
|||  
DB 34 PPPAY 38

Search completed: October 13, 2005, 15:09:26  
Job time : 14.4286 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 13.5165 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-16

Perfect score: 37

Sequence: 1 TTPPAY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	6	4	Aab83031 Human Sma
2	37	100.0	10	2	Aar98323 WBP-1 PY
3	37	100.0	14	2	Aaw38058 PPPPY mot
4	37	100.0	14	4	Aab83032 Human Sma
5	37	100.0	14	4	Aab83037 Human Sma
6	37	100.0	14	4	Aab83038 Human Sma
7	37	100.0	14	7	ADb49248 Biotinylia
8	37	100.0	16	2	Aaw82836 PY motif
9	37	100.0	16	2	Aaw82840 PY motif
10	37	100.0	16	2	Aaw82839 PY motif
11	37	100.0	216	3	Aab24036 Human PRO
12	37	100.0	216	3	Aab12136 Hydrophob
13	37	100.0	216	4	Aam93501 Human pol
14	37	100.0	216	5	Abg34048 Human Pro
15	37	100.0	216	6	ABO00496 Novel hum
16	37	100.0	216	6	ADA01306 Human PRO
17	37	100.0	216	6	ADA43735 Human sec
18	37	100.0	216	6	ADA43503 Human sec
19	37	100.0	216	6	ADA01178 Human PRO
20	37	100.0	216	7	ADA01062 Human sec
21	37	100.0	216	7	ADA43619 Human sec
22	37	100.0	216	7	ADA06881 Human PRO
23	37	100.0	216	7	ADA08369 Novel hum
24	37	100.0	216	7	ADB99662 Human PRO
25	37	100.0	216	7	ADB86945 Human PRO

26	37	100.0	216	7	ADB66100 Human sec
27	37	100.0	216	7	ADB99778 Human PRO
28	37	100.0	216	7	ADB99433 Novel hum
29	37	100.0	216	7	ADB5984 Human sec
30	37	100.0	216	7	ADC23382 Human tra
31	37	100.0	216	7	ADC26075 Human PRO
32	37	100.0	216	7	ADC26075 Human PRO
33	37	100.0	216	7	ADE04902 Human PRO
34	37	100.0	216	7	ADE11208 Human PRO
35	37	100.0	216	7	ADD88139 Human PRO
36	37	100.0	216	7	ADD95434 Human sec
37	37	100.0	216	7	ADE06364 Human PRO
38	37	100.0	216	7	ADE38139 Human PRO
39	37	100.0	216	7	ADD88255 Human PRO
40	37	100.0	216	7	ADD90836 Human sec
41	37	100.0	216	7	ADF99391 Human sec
42	37	100.0	216	7	ADG06484 Human PRO
43	37	100.0	216	7	ADG05435 Human PRO
44	37	100.0	216	7	ADG82436 Human PRO
45	37	100.0	216	8	ADE51689 Human sec
					ADE51805 Human sec

ALIGNMENTS

RESULT 1  
AAB83031  
ID AAB83031 standard; peptide; 6 AA.

XX AAB83031;  
XX 25-JUN-2001 (first entry)  
XX Human Smad 1 and Smad 5 PY motif consensus sequence.  
XX  
XX Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; PY motif; Smad 1; Smad 5.  
XX Homo sapiens.  
XX WO200116604-A1.  
XX 08-MAR-2001.  
XX 29-AUG-2000; 2000WO-US023729.  
XX 30-AUG-1999; 99US-00385918.  
XX (SIGN-) SIGNAL PHARM INC.  
XX Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
XX WPI; 2001-327913/34.  
XX Screening for modulators of TGF-beta and/or bone morphogenic protein  
XX (BMP) mediated signaling useful for treating cancer and osteoporosis by  
XX evaluating the ability of agents to modulate Smad protein degradation.  
XX Claim 4; Page 35; 75pp; English.  
XX The present sequence is the Smad PY motif consensus sequence. The PY  
XX motif binds to the WW domain of HECT (homologous to E6 carboxyl terminus)  
XX E3 ubiquitin ligase, resulting in ubiquitination of Smad by the E3  
XX ubiquitin ligase. The sequence is provided in a specification relating to  
XX a method for screening for agents that modulate transforming growth  
XX factor (TGF)-beta and/or bone morphogenic protein (BMP)-mediated  
XX signalling. The method involves evaluating the effect of an agent on  
XX binding of HECT E3 ubiquitin ligase WW domain to Smad PY motif, on  
XX ubiquitination of Smad protein by E3 ubiquitin ligase, or on the cellular

CC levels of Smad protein HECT E3 ubiquitin ligase activity. The method is  
 CC useful for stimulating bone formation in a patient or treating a  
 CC condition associated with insufficient TGF-beta and/or BMP-mediated cell  
 CC signalling. Agents that inhibit BMP-mediated signalling are useful for  
 CC treating inflammation, ageing, cancer and infectious diseases. Agents  
 CC that augment BMP-mediated signalling are useful for stimulating bone  
 CC anabolism as well as treating broken bones, osteoporosis, and acute or  
 CC chronic renal failure. Agents that inhibit TGF-mediated signalling are  
 CC useful for treating cancer, inflammation, neurodegeneration and fibrosis

XX Sequence 6 AA;

Query Match 100.0%; Score 37; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6  
 |||||  
 Db 1 TPPPAY 6

RESULT 2  
 AAR98323  
 ID AAR98323 standard; peptide; 10 AA.

XX AAR98323;

XX 30-AUG-1996 (first entry)

XX WBP-1 PY motif (P4A mutant).

XX WW domain; signal transduction; diagnosis; gene therapy;

XX Yes proto-oncogene associated protein; YAP; ligand; WBP-1; PY motif.

XX Synthetic.

XX WO9617061-A1.

XX 06-JUN-1996.

XX 30-NOV-1995; 95WO-US015512.

XX 01-DEC-1994; 94US-00348518.

XX 07-JUN-1995; 95US-00476509.

XX (UVRQ ) UNIV ROCKEFELLER.

XX (DELB-) DELBRUECK CENT MOLECULAR MEDICINE MAX.

XX Sudol M, Chen H, Bork P;

XX WPI, 1996-286829/29.

XX DNA encoding Yes proto-oncogene associated protein - used to modulate  
 PT intracellular signal transduction e.g. for treatment of muscular  
 PT dystrophy.

XX Example 5; Page 85; 126pp; English.

XX A series of mutants (AAR97698-700 and AAR98322-23) was made of the PY  
 CC motif (AAR97694) of WBP-1 (AAR97695), a ligand of the WW signalling  
 CC domain of a novel human proto-oncogene associated protein, YAP  
 CC (AAR97670). The mutant peptides, which also included the amino acids  
 CC flanking the PY motif, were expressed as GST fusion proteins in E. coli  
 CC SURE. Binding to the YAP WW domain was virtually abolished when proline  
 CC P2 (numbered according to position in the PY motif), P3 or Y5 was  
 CC substituted by alanine (AAR97698-700, respectively), and reduced approx.  
 CC 2-fold when P1 or P4 was altered to alanine (AAR98322-23, respectively)

XX Sequence 10 AA;

Query Match 100.0%; Score 37; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6  
 |||||  
 Db 2 TPPPAY 7

RESULT 3  
 AAW38058  
 ID AAW38058 standard; peptide; 14 AA.

XX AAW38058;

XX 23-APR-1998 (first entry)

XX PPPPY motif containing peptide used to bind WW domains.

XX Peptide recognition unit; WW domain; cell signalling; growth regulation;  
 KW cytoskeleton organisation; targeted drug screening; modulator;  
 KW WW domain interaction; YAP protein; dystrophin.

XX Synthetic.

XX WO9737223-A1.

XX 09-OCT-1997.

XX 03-APR-1997; 97WO-US005547.

XX 03-APR-1996; 96US-00630916.

XX (CYTO-) CYTOGEN CORP.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Pirozzi G, Kay BK, Fowlkes DM;

XX WPI; 1997-503234/46.

XX Identifying cell signalling and growth regulatory polypeptides by  
 PT reaction with multivalent recognition complex - polypeptides are useful  
 PT in targetted drug selection.

XX Disclosure; Fig 15A; 220pp; English.

XX Peptides AAW38057-67 contain ppppy-like motifs. The ppy motif is found  
 CC in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides  
 CC containing this residue have been shown to bind the YAP WW domain, but  
 CC not the WW domain from dystrophin or to a panel of SH3 domains. Peptides  
 CC AAW38057-67 were biotinylated and complexed with alkaline streptavidin,  
 CC and used in a cross affinity mapping experiment. They were tested for  
 CC their ability to bind to the 12 individual novel WW domains of WWP1  
 CC (AAW36794), WWP2 (AAW36795), WWP3 (AAW37696) and WWP4 (AAW36797), which  
 CC were expressed as glutathione-S-transferase expression proteins. The  
 CC present peptide is derived from WBP-1, and binds to some of the WW  
 CC domains of the novel proteins. The WW domain is a small functional  
 CC domain. Its name is derived from the observation that two tryptophan  
 CC residues, one in the amino terminal portion of the WW domain and one in  
 CC the carboxyl terminal portion, are conserved. Most proteins containing WW  
 CC domains have a function involving cell signalling and growth regulation  
 CC or the organisation of the cytoskeleton. Polypeptides containing a WW  
 CC domain are identified by treating a multivalent recognition unit complex  
 CC that has selective binding affinity for a WW domain, with many  
 CC polypeptides and identifying those with selective affinity for the  
 CC complex. Proteins containing WW domains are used for targeted drug  
 CC screening, i.e. to identify potential modulators of specific WW domain  
 CC interactions

XX Sequence 14 AA;

Query Match 100.0%; Score 37; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6

Db 3 TPPAY 8

RESULT 4  
AAB83032  
ID AAB83032 standard; peptide; 14 AA.  
XX  
AC AAB83032;  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Human Smad 1 and Smad 5 PY motif consensus sequence.  
XX  
KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; PY motif; Smad 5.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 10  
FT /label= Leu, Met  
XX  
PN WO200116604-A1.  
XX  
PD 08-MAR-2001.  
XX  
PF 29-AUG-2000; 2000WO-US023729.  
XX  
PR 30-AUG-1999; 99US-00385918.  
XX  
PA (SIGN-) SIGNAL PHARM INC.  
XX  
PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
XX WPI; 2001-327913/34.  
XX  
PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
PT evaluating the ability of agents to modulate Smad protein degradation.  
XX  
PS Claim 50; Page 41; 75pp; English.  
XX  
CC The present sequence is a Smad PY motif consensus sequence. The PY motif  
CC binds to the WW domain of HECT (homologous to E6 carboxyl terminus) E3  
CC ubiquitin ligase, resulting in ubiquitination of Smad by the E3 ubiquitin  
CC ligase. The sequence is provided in a specification relating to a method  
CC for screening for agents that modulate transforming growth factor (TGF)-  
CC beta and/or bone morphogenic protein (BMP)-mediated signalling. The  
CC method involves evaluating the effect of an agent on binding of HECT E3  
CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein  
CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
CC bone formation in a patient or treating a condition associated with  
CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
CC inhibit BMP-mediated signalling are useful for treating inflammation,  
CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
CC signalling are useful for stimulating bone anabolism as well as treating  
CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
CC that inhibit TGF-mediated signalling are useful for treating cancer,  
CC inflammation, neurodegeneration and fibrosis  
XX  
SQ Sequence 14 AA:  
Query Match 100.0%; Score 37; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TPPAY 6

Db 4 TPPAY 9

RESULT 5  
AAB83037  
ID AAB83037 standard; peptide; 14 AA.  
XX  
AC AAB83037;  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Human Smad PY motif #3.  
XX  
KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
KW antimicrobial; neuroprotective; transforming growth factor beta;  
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
KW renal failure; neurodegeneration; fibrosis; PY motif.  
XX  
OS Homo sapiens.  
XX  
PN WO200116604-A1.  
XX  
PD 08-MAR-2001.  
XX  
PF 29-AUG-2000; 2000WO-US023729.  
XX  
PR 30-AUG-1999; 99US-00385918.  
XX  
PA (SIGN-) SIGNAL PHARM INC.  
XX  
PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
XX WPI; 2001-327913/34.  
XX  
PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
PT evaluating the ability of agents to modulate Smad protein degradation.  
XX  
PS Example 1; Page 29; 75pp; English.  
XX  
CC The present sequence is a Smad PY motif. The PY motif binds to the WW  
CC domain of HECT (homologous to E6 carboxyl terminus) E3 ubiquitin ligase,  
CC resulting in ubiquitination of Smad by the E3 ubiquitin ligase. The  
CC sequence is provided in a specification relating to a method for  
CC screening for agents that modulate transforming growth factor (TGF)-beta  
CC and/or bone morphogenic protein (BMP)-mediated signalling. The method  
CC involves evaluating the effect of an agent on binding of HECT E3  
CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein  
CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
CC bone formation in a patient or treating a condition associated with  
CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
CC inhibit BMP-mediated signalling are useful for treating inflammation,  
CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
CC signalling are useful for stimulating bone anabolism as well as treating  
CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
CC that inhibit TGF-mediated signalling are useful for treating cancer,  
CC inflammation, neurodegeneration and fibrosis  
XX  
SQ Sequence 14 AA:  
Query Match 100.0%; Score 37; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TPPAY 6

```

RESULT 6
AAB83038
ID AAB83038 standard; peptide; 14 AA.
XX
AC AAB83038;
XX
DT 25-JUN-2001 (first entry)
XX
DE Human Smad PY motif #4.
XX
KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;
KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;
KW antimicrobial; neuroprotective; transforming growth factor beta;
KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;
KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;
KW renal failure; neurodegeneration; fibrosis; PY motif.
XX
OS Homo sapiens.
XX
PN WO200116604-A1.
XX
PD 08-MAR-2001.
XX
PF 29-AUG-2000; 2000WO-US023729.
XX
PR 30-AUG-1999; 99US-00385918.
XX
PA (SIGN-) SIGNAL PHARM INC.
XX
PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;
XX
DR WPI; 2001-327913/34.
XX
PT Screening for modulators of TGF-beta and/or bone morphogenic protein
PT (BMP) mediated signaling useful for treating cancer and osteoporosis by
PT evaluating the ability of agents to modulate Smad protein degradation.
XX
PS Example 2; Page 31; 75pp; English.
XX
CC The present sequence is a Smad PY motif. The PY motif binds to the WW
CC domain of HECT (homologous to E6 carboxyl terminus) E3 ubiquitin ligase,
CC resulting in ubiquitination of Smad by the E3 ubiquitin ligase. The
CC sequence is provided in a specification relating to a method for
CC screening for agents that modulate transforming growth factor (TGF)-beta
CC and/or bone morphogenic protein (BMP)-mediated signalling. The method
CC involves evaluating the effect of an agent on binding of HECT E3
CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad
CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein
CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating
CC bone formation in a patient or treating a condition associated with
CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that
CC inhibit BMP-mediated signalling are useful for treating inflammation,
CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated
CC signalling are useful for stimulating bone anabolism as well as treating
CC broken bones, osteoporosis, and acute or chronic renal failure. Agents
CC that inhibit TGF-mediated signalling are useful for treating cancer,
CC inflammation, neurodegeneration and fibrosis
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 37; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPPPAY 6
DB 4 TPPPAY 9
RESULT 7
ADB49248
ID ADB49248 standard; peptide; 14 AA.
XX
AC ADB49248;
XX
DT 04-DEC-2003 (first entry)
XX
DE Biotinylated WW domain binding peptide #31.
XX
KW WW domain; drug candidate screening; drug discovery; drug modification;
KW drug refinement; immunogen; WW binding protein; WW domain.
XX
OS Unidentified.
XX
PN US2003077577-A1.
XX
PD 24-APR-2003.
XX
PF 28-JUN-2002; 2002US-00185050.
XX
PR 03-APR-1996; 96US-00630916.
PR 03-APR-1997; 97US-00826516.
XX
PA (PIROV/) PIROZZI G.
PA (KAYB/) KAY B K.
PA (FOWL/) FOWLKES D M.
XX
PI Pirozzi G, Kay BK, Fowlkes DM;
XX
DR WPI; 2003-635075/60.
XX
PT Novel purified polypeptide comprising WW domain, useful for drug
PT discovery, modification and refinement, for discovering polypeptides
PT involved in pharmacological activities, or as an immunogen to generate
PT antibodies.
XX
PS Example; Fig 15A; 133pp; English.
XX
CC The invention describes a purified polypeptide (I) comprising a WW domain
CC which has a sequence (S1) selected from 11 sequences fully defined in the
CC specification, a sequence (S2) selected from 48 sequences fully defined
CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
CC amino acids fully defined in the specification. (I) is useful for
CC screening a potential drug candidate, by allowing (I) to come into
CC contact with at least one recognition unit having a selective affinity
CC for the WW domain in (I), in the presence of an amount of a potential
CC drug candidate, such that (I) and the recognition unit are capable of
CC interacting when brought into contact with one another in the absence of
CC the drug candidate, and determining the effect, if any, of the presence
CC of the amount of the drug candidate on the interaction of (I) with the
CC recognition unit. (I) is useful for drug discovery, modification and
CC refinement, for discovering polypeptides involved in pharmacological
CC activities, or as an immunogen to generate antibodies. This is the amino
CC acid sequence of a WW domain binding peptide.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 37; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPPPAY 6
DB 3 TPPPAY 8
RESULT 8
AAW82836
ID AAW82836 standard; peptide; 16 AA.
XX
AC AAW82836;
XX
DT 01-FEB-1999 (first entry)
XX
PY PY motif of the human MAD-1 protein.
XX

```

KW Endothelial MAD interactor protein 1; mothers against dpp; MAD-1;  
 KW decapentaplegic; human; transforming growth factor-beta; TGF-beta;  
 KW EMII deficiency; cardiovascular disease; atherosclerosis; ischaemia;  
 KW reperfusion; hypertension; restenosis; arterial inflammation;  
 KW proliferative disease; cancer; colorectal; pancreatic; antisense;  
 KW angiogenesis; autoimmune disease; fibrosis; wound healing.  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO9845467-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US007356.  
 XX  
 PR 10-APR-1997; 97US-00844312.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Gimeno CJ, Falb DA;  
 XX  
 DR WPI; 1998-583204/49.  
 XX  
 PT Nucleic acid encoding endothelial MAD interactor I - for modulating cell  
 PT proliferation and differentiation, e.g. in cases of atherosclerosis and  
 PT cancer of colon or pancreas.  
 XX  
 PS Example 4; Page 67; 95pp; English.  
 XX  
 CC The present sequence represents the PY motif of a MAD-1 (mothers against  
 CC dpp (decapentaplegic)) protein. The specification describes a human  
 CC endothelial MAD interactor 1 protein EMII. The protein modulates  
 CC transforming growth factor-beta (TGF-beta) response in TGF beta  
 CC responsive cells. The EMII protein is used to treat EMII deficiency. They  
 CC are especially used for treatment of cardiovascular disease (specifically  
 CC atherosclerosis but also ischaemia/reperfusion, hypertension, restenosis  
 CC and arterial inflammation) or proliferative diseases, especially cancer  
 CC of gut-derived cells, e.g. colorectal or pancreatic cancer, also to  
 CC modulate angiogenesis, treat autoimmune disease or fibrosis and to  
 CC regulate wound healing  
 XX  
 SQ Sequence 16 AA;  
 Query Match 100.0%; Score 37; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db |||||  
 7 TPPPAY 12  
 RESULT 9  
 AAW82840  
 ID AAW82840 standard; peptide; 16 AA.  
 XX  
 AC AAW82840;  
 XX  
 DT 01-FEB-1999 (first entry)  
 XX  
 DE PY motif of the Drosophila MAD protein.  
 XX  
 KW Endothelial MAD interactor protein 1; mothers against dpp; MAD;  
 KW decapentaplegic; human; transforming growth factor-beta; TGF-beta;  
 KW EMII deficiency; cardiovascular disease; atherosclerosis; ischaemia;  
 KW reperfusion; hypertension; restenosis; arterial inflammation;  
 KW proliferative disease; cancer; colorectal; pancreatic; antisense;  
 KW angiogenesis; autoimmune disease; fibrosis; wound healing.  
 XX  
 OS Drosophila sp.  
 XX  
 PN WO9845467-A1.  
 XX

PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US007356.  
 XX  
 PR 10-APR-1997; 97US-00844312.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Gimeno CJ, Falb DA;  
 XX  
 DR WPI; 1998-583204/49.  
 XX  
 PT Nucleic acid encoding endothelial MAD interactor I - for modulating cell  
 PT proliferation and differentiation, e.g. in cases of atherosclerosis and  
 PT cancer of colon or pancreas.  
 XX  
 PS Example 4; Page 67; 95pp; English.  
 XX  
 CC The present sequence represents the PY motif of a MAD (mothers against  
 CC dpp (decapentaplegic)) protein. The specification describes a human  
 CC endothelial MAD interactor 1 protein EMII. The protein modulates  
 CC transforming growth factor-beta (TGF-beta) response in TGF beta  
 CC responsive cells. The EMII protein is used to treat EMII deficiency. They  
 CC are especially used for treatment of cardiovascular disease (specifically  
 CC atherosclerosis but also ischaemia/reperfusion, hypertension, restenosis  
 CC and arterial inflammation) or proliferative diseases, especially cancer  
 CC of gut-derived cells, e.g. colorectal or pancreatic cancer, also to  
 CC modulate angiogenesis, treat autoimmune disease or fibrosis and to  
 CC regulate wound healing  
 XX  
 SQ Sequence 16 AA;  
 Query Match 100.0%; Score 37; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db |||||  
 7 TPPPAY 12  
 RESULT 10  
 AAW82839  
 ID AAW82839 standard; peptide; 16 AA.  
 XX  
 AC AAW82839;  
 XX  
 DT 01-FEB-1999 (first entry)  
 XX  
 DE PY motif of the Smad-5 protein.  
 XX  
 KW Endothelial MAD interactor protein 1; mothers against dpp; Smad-5;  
 KW decapentaplegic; human; transforming growth factor-beta; TGF-beta;  
 KW EMII deficiency; cardiovascular disease; atherosclerosis; ischaemia;  
 KW reperfusion; hypertension; restenosis; arterial inflammation;  
 KW proliferative disease; cancer; colorectal; pancreatic; antisense;  
 KW angiogenesis; autoimmune disease; fibrosis; wound healing.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9845467-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US007356.  
 XX  
 PR 10-APR-1997; 97US-00844312.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Gimeno CJ, Falb DA;  
 XX  
 DR WPI; 1998-583204/49.  
 XX





CC cells. Membrane proteins have important roles as signal receptors, ion  
CC channels and transporters. The present sequence is a human protein which  
CC has at least one hydrophobic domain. This protein may be a secretory or a  
CC membrane protein. The present protein may have cytokine and cell  
CC proliferation/differentiation activity, immune stimulating or suppressing  
CC activity, haematopoiesis activity, tissue growth activity,  
CC actin/inhibin activity, chemotactic/chemokinetic activity,  
CC and thrombolytic activity, anti-inflammatory activity and tumour  
CC inhibition activity. The present protein could therefore be used for  
CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
CC disease, and cancer

XX SQ Sequence 216 AA;

Query Match 100.0%; Score 37; DB 3; Length 216;  
Best Local Similarity 100.0%; Pred. NO. 3.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPPPAY 6  
Db 190 TPPPAY 195

RESULT 13  
AAM93501  
ID AAM93501 standard; protein; 216 AA.

XX AC AAM93501;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3209.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-00114089.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI; 2001-524255/58.

XX DR N-PSDB; AAK94429.

XX PT 830 Primers useful for synthesizing full length cDNA clones and their use

XX PT in Genetic manipulation.

XX PS Claim 8; SEQ ID NO 3209; 1380pp + Sequence Listing; English.

XX CC The invention relates to primers for synthesising full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been isolated  
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
CC been determined. Primers for synthesising the full length cDNA are useful  
CC for clarifying the function of the protein encoded by the cDNA. The full  
CC length clones were obtained by construction of full length enriched cDNA  
CC libraries that were synthesised by the oligo-capping method. The primers  
CC enable the production of the full length cDNA easily without any special  
CC methods. The present sequence is a polypeptide encoded by a full length  
CC human cDNA of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in CD-ROM  
CC format directly from EPO

XX

SQ Sequence 216 AA;

Query Match 100.0%; Score 37; DB 4; Length 216;  
Best Local Similarity 100.0%; Pred. NO. 3.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPPPAY 6  
Db 190 TPPPAY 195

RESULT 14  
ABG34048  
ID ABG34048 standard; protein; 216 AA.

XX AC ABG34048;

XX DT 15-JUL-2002 (first entry)

XX DE Human Pro peptide #19.

XX KW Human; PRO; secreted protein; transmembrane protein; genetic disorder;  
XX KW tumour; cancer.

XX OS Homo sapiens.

XX PN WO200224888-A2.

XX PD 28-MAR-2002.

XX PF 29-AUG-2001; 2001WO-US027099.

XX PR 01-SEP-2000; 2000US-0229896P.

XX PR 05-SEP-2000; 2000US-0230621P.

XX PR 22-SEP-2000; 2000US-0235147P.

XX PR 10-NOV-2000; 2000WO-US030873.

XX PR 12-JAN-2001; 2001US-0261878P.

XX PR 16-JAN-2001; 2001US-0261910P.

XX PR 16-JAN-2001; 2001US-0261939P.

XX PR 25-JAN-2001; 2001US-0264395P.

XX PR 02-FEB-2001; 2001US-0266421P.

XX PR 09-FEB-2001; 2001US-0267623P.

XX PR 28-FEB-2001; 2001WO-US006520.

XX PR 09-MAR-2001; 2001US-0274399P.

XX PR 03-APR-2001; 2001US-0280982P.

XX PR 04-APR-2001; 2001US-0282129P.

XX PR 09-MAY-2001; 2001US-0290589P.

XX PR 25-MAY-2001; 2001WO-US017092.

XX PR 01-JUN-2001; 2001WO-US017800.

XX PR 20-JUN-2001; 2001WO-US019692.

XX PR 29-JUN-2001; 2001WO-US021066.

XX PR 09-JUL-2001; 2001WO-US021735.

XX PA (GETH ) GENENTECH INC.

XX PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;

XX PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;

XX PI Fong S;

XX DR WPI; 2002-362426/39.

XX DR N-PSDB; ABK69979.

XX PT New PRO polypeptides and polynucleotides encoding the polypeptides,  
XX useful in gene therapy, chromosome identification, tissue typing, or for  
XX genetic analysis of individuals with genetic disorders.

XX PS Claim 11; Fig 38; 218pp; English.

XX CC This invention relates to the cDNA and protein sequences of novel  
XX secreted and transmembrane polypeptides PRO polypeptides. The invention  
XX also comprises a method for producing the proteins of the invention by

CC recombinant means and antibodies specific for the protein of the  
 CC invention. The antibody may be used for detecting the PRO proteins of the  
 CC invention and may be used to modify their activity. polynucleotides may  
 CC be used as hybridisation probes for a cDNA library to isolate the full-  
 CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation  
 CC probes for mapping the gene which encodes that PRO and for genetic  
 CC analysis of individuals with genetic disorders, in assays to identify  
 CC other proteins or molecules involved in binding reaction, to generate  
 CC transgenic animals or knock-out animals which in turn are useful in the  
 CC development and screening of therapeutically useful reagents, for  
 CC chromosome identification, and tissue typing. The PRO polypeptides are  
 CC useful in gene therapy, and as molecular weight markers for protein  
 CC electrophoresis purposes. The sequences may also be used to detect  
 CC overexpression on PRO polypeptides in cancerous tumours and for screening  
 CC for differentially expressed genes using microarray technology. The  
 CC present sequence represents a human PRO protein of the invention  
 XX

XX Sequence 216 AA;

Query Match 100.0%; Score 37; DB 5; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 190 TPPPAY 195

RESULT 15

ABO00496

ID ABO00496 standard; protein; 216 AA.

XX AC ABO00496;

DT 06-AUG-2003 (first entry)

DE Novel human polypeptide #83.

XX Human; angiogenesis; cytokine; cell proliferation; pluripotent;  
 KW cell differentiation; totipotent; stem cell; transplantation; bio-sensor;  
 KW neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;  
 KW nerve; brain tissue; central nervous system disease;  
 KW myeloid disorder; lymphoid cell disorder; haematopoiesis; bone;  
 KW regeneration; cartilage; tendon; ligament; nerve tissue growth;  
 KW tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;  
 KW osteoarthritis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung fibrosis; liver fibrosis; reperfusion injury;  
 KW immune deficiency; infection; autoimmune disorder; allergic reaction;  
 KW thrombolytic; thrombosis; coagulation disorder; hereditary disorder;  
 KW biorhythm; circadian cycle; fertility; metabolism; catabolism; anabolism;  
 KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;  
 KW haemostatic; vulnery; antiulcer; osteopathic; antiarthritic;  
 KW vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;  
 KW anti rheumatic; antidiabetic; antiasthmatic; cytostatic; virucide.

XX Homo sapiens.

XX WO2003023013-A2.

XX 20-MAR-2003.

XX 13-SEP-2002; 2002WO-US029001.

XX 13-SEP-2001; 2001US-0322511P.

PR 12-SEP-2002; 2002US-00243552.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX WPI; 2003-313249/30.

DR N-PSDB; ACD05573.

XX Novel nucleic acids and polypeptides for diagnosis, treatment of central  
 PT and peripheral nervous system diseases and neuropathies, such as  
 PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX lateral sclerosis.

PS Claim 20; SEQ ID NO 419; 300pp; English.

XX The present invention relates to the isolation of novel human  
 CC polynucleotide sequences and their encoding polypeptides. The novel  
 CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell  
 CC proliferation, cell differentiation, antiinflammatory, and stem cell  
 CC growth factor activities. The polypeptides are involved in the  
 CC proliferation, differentiation and survival of pluripotent and totipotent  
 CC stem cells, and are useful for re-engineering damaged or diseased  
 CC tissues, transplantation, manufacture of bio-pharmaceuticals and  
 CC development of bio-sensors. The polypeptides can be used to manipulate  
 CC stem cells in culture to give rise to neuroepithelial cells that can be  
 CC used to augment or replace cells damaged by illness, autoimmune disease,  
 CC accidental damage or genetic disorders. The polypeptides induce the  
 CC proliferation of neural cells and regeneration of nerve and brain tissue  
 CC and are useful for the treatment of central and peripheral nervous system  
 CC diseases and neuropathies, such as Alzheimer's, Parkinson's disease,  
 CC Huntington's disease, amyotrophic lateral sclerosis (ALS). The  
 CC polypeptides are also involved in chemotactic or chemokinetic activity,  
 CC regulation of haematopoiesis and are useful for treating myeloid or  
 CC lymphoid cell disorders, platelet disorders such as thrombocytopaenia and  
 CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
 CC growth, in tissue repair, healing of burns, incisions, ulcers, for  
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, and  
 CC periodontal disease. The polypeptides are also useful for gut protection  
 CC or regeneration and treatment of lung or liver fibrosis, reperfusion  
 CC injury in various tissues, various immune deficiencies and disorders  
 CC including severe combined immunodeficiency (SCID), bacterial or fungal  
 CC infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid  
 CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and  
 CC conditions, such as asthma or other respiratory problems. The  
 CC polypeptides are involved in thrombolysis or thrombosis and are useful in  
 CC treatment of various coagulation disorders (including hereditary  
 CC disorders such as haemophilia) or to enhance coagulation and other  
 CC haemostatic events in treating wounds resulting from trauma, surgery or  
 CC other causes. The polypeptides exhibit immune stimulating or immune  
 CC suppressing activity, and are useful for treating autoimmune diseases or  
 CC cancer. They also inhibit the growth, infection or function of infectious  
 CC agents such as bacteria, fungi, viruses, effect biorhythms or circadian  
 CC cycles of rhythms, fertility of male or female subjects, metabolism,  
 CC catabolism, and anabolism. ABO00414-ABO00749 represent the novel  
 CC polypeptides of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 216 AA;

Query Match 100.0%; Score 37; DB 6; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 190 TPPPAY 195

Search completed: October 13, 2005, 15:00:36  
 Job time : 15.5165 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 2.85714 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-16  
Perfect score: 37  
Sequence: 1 TPPPAY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	10	3 US-08-476-509B-35	Sequence 35, Appl
2	37	100.0	14	3 US-08-630-918A-54	Sequence 54, Appl
3	37	100.0	16	2 US-08-844-312-12	Sequence 12, Appl
4	37	100.0	16	2 US-08-844-312-15	Sequence 15, Appl
5	37	100.0	16	2 US-08-844-312-16	Sequence 16, Appl
6	37	100.0	128	4 US-09-502-540-10130	Sequence 10130, A
7	37	100.0	403	4 US-09-949-016-10865	Sequence 10865, A
8	37	100.0	455	3 US-08-840-767-43	Sequence 43, Appl
9	37	100.0	456	3 US-08-840-767-11	Sequence 11, Appl
10	37	100.0	464	4 US-08-580-031A-14	Sequence 14, Appl
11	37	100.0	465	3 US-08-701-582D-13	Sequence 13, Appl
12	37	100.0	465	3 US-09-063-893A-19	Sequence 19, Appl
13	37	100.0	465	3 US-08-840-767-8	Sequence 8, Appl
14	37	100.0	465	3 US-08-840-767-48	Sequence 48, Appl
15	37	100.0	465	3 US-08-840-767-52	Sequence 52, Appl
16	37	100.0	465	3 US-09-096-776B-6	Sequence 6, Appl
17	37	100.0	465	3 US-09-096-776B-10	Sequence 10, Appl
18	37	100.0	465	4 US-09-923-922-6	Sequence 6, Appl
19	37	100.0	465	4 US-09-923-922-10	Sequence 10, Appl
20	37	100.0	475	3 US-08-840-767-10	Sequence 10, Appl
21	37	100.0	492	4 US-09-949-016-10447	Sequence 10447, A
22	34	91.9	77	4 US-09-513-999C-6393	Sequence 6393, Ap
23	34	91.9	219	1 US-08-843-993-4	Sequence 4, Appl
24	34	91.9	219	3 US-09-059-520A-4	Sequence 4, Appl
25	34	91.9	219	3 US-09-334-275-4	Sequence 4, Appl
26	34	91.9	219	4 US-09-519-172-74	Sequence 74, Appl
27	34	91.9	390	4 US-09-198-452A-254	Sequence 254, Appl

28	34	91.9	390	4	US-09-438-185A-243	Sequence 243, App
29	34	91.9	444	4	US-09-354-221-4	Sequence 4, Appl
30	34	91.9	530	4	US-09-949-016-9519	Sequence 9519, Ap
31	34	91.9	648	2	US-08-817-436A-2	Sequence 2, Appl
32	33	89.2	16	2	US-08-844-312-13	Sequence 13, Appl
33	33	89.2	16	2	US-08-844-312-14	Sequence 14, Appl
34	33	89.2	152	2	US-08-752-844-4	Sequence 4, Appl
35	33	89.2	152	2	US-08-591-196-4	Sequence 4, Appl
36	33	89.2	152	3	US-09-192-838B-4	Sequence 4, Appl
37	33	89.2	152	4	US-09-293-533-4	Sequence 4, Appl
38	33	89.2	152	4	US-09-324-191-4	Sequence 4, Appl
39	33	89.2	153	3	US-09-096-244-4	Sequence 4, Appl
40	33	89.2	169	4	US-09-270-767-37020	Sequence 37020, A
41	33	89.2	169	4	US-09-270-767-52237	Sequence 52237, A
42	33	89.2	401	1	US-08-368-803-7	Sequence 7, Appl
43	33	89.2	422	2	US-08-663-566A-5	Sequence 5, Appl
44	33	89.2	422	2	US-08-023-610-5	Sequence 5, Appl
45	33	89.2	422	2	US-08-288-065A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1  
US-08-476-509B-35  
; Sequence 35, Application US/08476509B  
; Patent No. 6034212  
; GENERAL INFORMATION:  
; APPLICANT: SUDOL, MARIUS  
; APPLICANT: PEER, BORK  
; APPLICANT: HENRY, CHEN  
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,509B  
; FILING DATE: 01-DEC-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-476-509B-35

Query Match 100.0%; Score 37; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.7; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;



```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-312-15

Query Match      100.0%; Score 37; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTPPAY 6
Db 7 TTPPAY 12

RESULT 5
US-08-844-312-16
; Sequence 16, Application US/08844312
; Patent No. 5948639
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639el TGF-{SYMBOL 98 \f "Symbol"} Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,312
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-312-16

Query Match      100.0%; Score 37; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTPPAY 6
Db 7 TTPPAY 12

RESULT 6
US-09-902-540-10130
; Sequence 10130, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
```

```
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10130
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10130

Query Match      100.0%; Score 37; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTPPAY 6
Db 123 TTPPAY 128

RESULT 7
US-09-949-016-10865
; Sequence 10865, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10865
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10865

Query Match      100.0%; Score 37; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTPPAY 6
Db 236 TTPPAY 241

RESULT 8
US-08-840-767-43
; Sequence 43, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggins, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 455
; TYPE: PRT
; ORGANISM: D. melanogaster
US-08-840-767-43

Query Match          100.0%; Score 37; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 220 TPPPAY 225

RESULT 9
US-08-840-767-11
; Sequence 11, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggs, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-08-840-767-11

Query Match          100.0%; Score 37; DB 3; Length 456;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 220 TPPPAY 225

RESULT 10
US-08-580-031A-14
; Sequence 14, Application US/08580031A
; Patent No. 6428977
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathan M.
; APPLICANT: Woolf, Tod M.
; APPLICANT: Jin, Ping
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: The "signalin" Family of TGFb Signal
; TITLE OF INVENTION: Transduction Proteins, and Uses Related Thereto
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/580,031A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: ONV-019.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)832-1000
; TELEFAX: (617)832-7000
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-580-031A-14

Query Match          100.0%; Score 37; DB 4; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 221 TPPPAY 226

RESULT 11
US-08-701-582D-13
; Sequence 13, Application US/08701582D
; Patent No. 6017755
; GENERAL INFORMATION:
; APPLICANT: WRANA, Jeffrey
; APPLICANT: ATTISANO, Lilliana
; APPLICANT: SCHERER, Stephen W.
; TITLE OF INVENTION: MADR2 TUMOUR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,582D
; FILING DATE: 22-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 024916-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-582D-13

Query Match          100.0%; Score 37; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 TTPPAY 6
Db      222 TTPPAY 227

RESULT 12
US-09-063-893A-19
; Sequence 19, Application US/09063893A
; Patent No. 620464
; GENERAL INFORMATION:
; APPLICANT: Kenji OKAZAKI et al.
; TITLE OF INVENTION: A NOVEL SIGNAL TRANSDUCTION FACTOR AND A
; TITLE OF INVENTION: GENE ENCODING THE SAME
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,893A
; FILING DATE: April 22, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: Swiss-Webster/NIH
US-09-063-893A-19

Query Match      100.0%; Score 37; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTPPAY 6
Db      222 TTPPAY 227

RESULT 13
US-08-840-767-8
; Sequence 8, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggins, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548

```

```

; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-840-767-8

Query Match      100.0%; Score 37; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTPPAY 6
Db      221 TTPPAY 226

RESULT 14
US-08-840-767-48
; Sequence 48, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggins, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-840-767-48

Query Match      100.0%; Score 37; DB 3; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTPPAY 6
Db      221 TTPPAY 226

RESULT 15
US-08-840-767-52
; Sequence 52, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggins, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 465
; TYPE: PRT

```

! ORGANISM: Homo sapiens  
US-08-840-767-52

Query Match 100.0%; Score 37; DB 3; Length 465;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTPPAY 6  
|||  
Db 221 TTPPAY 226

Search completed: October 13, 2005, 14:03:07  
Job time : 3.98214 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 12.7912 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-16

Perfect score: 37

Sequence: 1 TPPPAY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	6	14 US-10-307-956-16	Sequence 16, Appl
2	37	100.0	14	14 US-10-185-050-54	Sequence 54, Appl
3	37	100.0	14	14 US-10-307-956-17	Sequence 17, Appl
4	37	100.0	14	14 US-10-307-956-22	Sequence 22, Appl
5	37	100.0	14	14 US-10-307-956-23	Sequence 23, Appl
6	37	100.0	110	15 US-10-424-599-283015	Sequence 283015,
7	37	100.0	132	16 US-10-767-701-48923	Sequence 48923, A
8	37	100.0	177	16 US-10-767-701-31853	Sequence 31853, A
9	37	100.0	195	16 US-10-425-115-359021	Sequence 359021,
10	37	100.0	216	14 US-10-245-752-38	Sequence 38, Appl
11	37	100.0	216	14 US-10-245-859-38	Sequence 38, Appl

12	37	100.0	216	14	US-10-245-103-38	Sequence 38, Appl
13	37	100.0	216	14	US-10-245-107-38	Sequence 38, Appl
14	37	100.0	216	14	US-10-245-143-38	Sequence 38, Appl
15	37	100.0	216	14	US-10-245-771-38	Sequence 38, Appl
16	37	100.0	216	14	US-10-245-851-38	Sequence 38, Appl
17	37	100.0	216	14	US-10-245-883-38	Sequence 38, Appl
18	37	100.0	216	14	US-10-237-535-38	Sequence 38, Appl
19	37	100.0	216	14	US-10-238-183-38	Sequence 38, Appl
20	37	100.0	216	14	US-10-238-283-38	Sequence 38, Appl
21	37	100.0	216	14	US-10-238-370-38	Sequence 38, Appl
22	37	100.0	216	14	US-10-245-055-38	Sequence 38, Appl
23	37	100.0	216	14	US-10-245-147-38	Sequence 38, Appl
24	37	100.0	216	14	US-10-245-730-38	Sequence 38, Appl
25	37	100.0	216	14	US-10-245-739-38	Sequence 38, Appl
26	37	100.0	216	14	US-10-246-210-38	Sequence 38, Appl
27	37	100.0	216	14	US-10-239-196-38	Sequence 38, Appl
28	37	100.0	216	14	US-10-243-024-38	Sequence 38, Appl
29	37	100.0	216	14	US-10-243-409-38	Sequence 38, Appl
30	37	100.0	216	14	US-10-245-621-38	Sequence 38, Appl
31	37	100.0	216	14	US-10-245-880-38	Sequence 38, Appl
32	37	100.0	216	14	US-10-245-033-38	Sequence 38, Appl
33	37	100.0	216	14	US-10-243-095-38	Sequence 38, Appl
34	37	100.0	216	14	US-10-245-185-38	Sequence 38, Appl
35	37	100.0	216	14	US-10-245-427-38	Sequence 38, Appl
36	37	100.0	216	14	US-10-245-473-38	Sequence 38, Appl
37	37	100.0	216	14	US-10-245-770-38	Sequence 38, Appl
38	37	100.0	216	14	US-10-245-877-38	Sequence 38, Appl
39	37	100.0	216	14	US-10-246-976-38	Sequence 38, Appl
40	37	100.0	216	14	US-10-243-320-38	Sequence 38, Appl
41	37	100.0	216	14	US-10-242-743-38	Sequence 38, Appl
42	37	100.0	216	14	US-10-242-845-38	Sequence 38, Appl
43	37	100.0	216	14	US-10-237-636-38	Sequence 38, Appl
44	37	100.0	216	14	US-10-238-325-38	Sequence 38, Appl
45	37	100.0	216	14	US-10-238-346-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1  
US-10-307-956-16  
; Sequence 16, Application US/10307956  
; Publication No. US20030119072A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoechst, Merl F.  
; APPLICANT: Xie, Weilin  
; APPLICANT: Murray, Brion  
; APPLICANT: Mercurio, Frank  
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL  
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS  
; FILE REFERENCE: 860098.433  
; CURRENT APPLICATION NUMBER: US/10/307,956  
; CURRENT FILING DATE: 2002-12-02  
; PRIOR APPLICATION NUMBER: US/09/385,918  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-307-956-16

Query Match 100.0%; Score 37; DB 14; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 1 TPPPAY 6

RESULT 2

```
US-10-185-050-54
; Sequence 54, Application US/10185050
; Publication No. US2003007577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/10/185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-185-050-54

Query Match 100.0%; Score 37; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 3 TPPPAY 8

RESULT 3
; Sequence 17, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-23

Query Match 100.0%; Score 37; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 4 TPPPAY 9

RESULT 5
; Sequence 23, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 14
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (10)...(10)
; OTHER INFORMATION: Xaa = Leucine or Methionine
US-10-307-956-17

Query Match 100.0%; Score 37; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 4 TPPPAY 9

RESULT 4
; Sequence 22, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-22

Query Match 100.0%; Score 37; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 4 TPPPAY 9

RESULT 5
; Sequence 23, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brian
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 14
```

```
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-23

Query Match      100.0%; Score 37; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db 4 TPPPAY 9

RESULT 6
US-10-424-598-283015
; Sequence 283015, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283015
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97585C.1.pep
US-10-424-599-283015

Query Match      100.0%; Score 37; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db 29 TPPPAY 34

RESULT 7
US-10-767-701-48923
; Sequence 48923, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48923
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-020-PI-K1-G7.pep
US-10-767-701-48923

Query Match      100.0%; Score 37; DB 16; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db 125 TPPPAY 130

RESULT 8
US-10-767-701-31853
; Sequence 31853, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 31853
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(177)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C107513_1.pep
US-10-767-701-31853

Query Match      100.0%; Score 37; DB 16; Length 177;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db 4 TPPPAY 9

RESULT 9
US-10-425-115-359021
; Sequence 359021, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 359021
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(195)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_90594C.1.pep
US-10-425-115-359021

Query Match      100.0%; Score 37; DB 16; Length 195;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db 125 TPPPAY 130
```

```
RESULT 10
US-10-245-752-38
; Sequence 38, Application US/10245752
; Publication No. US20030064473A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C66
; CURRENT APPLICATION NUMBER: US/10/245,752
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-752-38
```

```
Query Match 100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TTPPAY 6
Db 190 TTPPAY 195
```

```
RESULT 11
US-10-245-859-38
; Sequence 38, Application US/10245859
; Publication No. US20030064474A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
```

```
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C78
; CURRENT APPLICATION NUMBER: US/10/245,859
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-859-38
```

```
Query Match 100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TTPPAY 6
Db 190 TTPPAY 195
```

```
RESULT 12
US-10-245-103-38
; Sequence 38, Application US/10245103
; Publication No. US20030068778A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C112
; CURRENT APPLICATION NUMBER: US/10/245,103
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
```

```
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-103-38
```

```
Query Match 100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 1 TPPPAY 6
DB 190 TPPPAY 195
```

```
RESULT 13
US-10-245-107-38
; Sequence 38, Application US/10245107
; Publication No. US20030068779A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C71
; CURRENT APPLICATION NUMBER: US/10/245,107
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C71
; CURRENT APPLICATION NUMBER: US/10/245,107
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-103-38
```

```
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-107-38
```

```
Query Match 100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPPPAY 6
DB 190 TPPPAY 195
```

```
RESULT 14
US-10-245-143-38
; Sequence 38, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-143-38
```

```
Query Match 100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPPPAY 6
```

```
Db.      190 TTPPAY 195
|||||
RESULT 15
US-10-245-771-38
; Sequence 38, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC98
; CURRENT APPLICATION NUMBER: US/10/245,771
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 38
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-771-38
Query Match      100.0%; Score 37; DB 14; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 TTPPAY 6
Db      190 TTPPAY 195
|||||
Search completed: October 13, 2005, 14:50:03
Job time : 13.7912 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:36:01 ; Search time 2.50549 Seconds  
(without alignments)  
230.414 Million cell updates/sec

Title: US-09-385-918-16

Perfect score: 37

Sequence: 1 TPPPAY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79: \*  
1: Pirl: \*  
2: pirl2: \*  
3: pirl3: \*  
4: pirl4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	367	2 T24298	hypothetical prote
2	37	100.0	455	2 S55019	MAD polypeptide -
3	37	100.0	465	2 S68987	transcription acti
4	37	100.0	520	2 D84471	probable membrane
5	37	100.0	556	2 A10134	probable phage pro
6	37	100.0	632	2 I51682	epithelial sodium
7	37	100.0	1095	2 T20528	hypothetical prote
8	34	91.9	82	2 T41925	hypothetical prote
9	34	91.9	88	2 T44170	hypothetical prote
10	34	91.9	88	2 T43984	hypothetical prote
11	34	91.9	93	2 T09318	BolF1 protein - hu
12	34	91.9	219	2 S35643	BRE2 protein - hu
13	34	91.9	384	2 A86521	hypothetical prote
14	34	91.9	384	2 E72100	hypothetical prote
15	34	91.9	432	2 A28913	regulatory protein
16	34	91.9	483	2 T24856	hypothetical prote
17	34	91.9	497	1 S33938	penton protein (li
18	34	91.9	520	2 G88946	protein T12A7.2 (i
19	34	91.9	908	2 G82824	aconitase XF0290 (
20	33	89.2	99	2 T33486	hypothetical prote
21	33	89.2	214	2 D84808	hypothetical prote
22	33	89.2	240	2 T15785	hypothetical prote
23	33	89.2	267	2 T20096	hypothetical prote
24	33	89.2	274	2 D71443	hypothetical prote
25	33	89.2	285	2 T20506	hypothetical prote
26	33	89.2	296	2 JC7267	paired-type homeod
27	33	89.2	357	2 B81965	hemoglobin-haptogl
28	33	89.2	379	2 T16213	APX-1 protein homo
29	33	89.2	424	2 S71798	MAD-3 protein homo

30	33	89.2	444	1 F59904	adenosylmethionine
31	33	89.2	467	2 S71797	MAD-2 protein homo
32	33	89.2	484	2 JE0341	R-SWAD protein, Sm
33	33	89.2	513	2 D88991	protein apx-1 [imp
34	33	89.2	599	2 T10798	pherophorin-S - Vo
35	33	89.2	603	2 T14145	NADH2 dehydrogenas
36	33	89.2	660	2 D87331	TonB-dependent rec
37	33	89.2	665	2 H87468	ubiquinol oxidase
38	33	89.2	668	2 B54759	ba-type ubiquinol
39	33	89.2	751	2 T15230	hypothetical prote
40	33	89.2	766	2 T03218	armadillo-like pro
41	33	89.2	887	2 T11566	envelope glycoprot
42	33	89.2	1101	2 S15271	endoglucanase C (E
43	33	89.2	1227	2 T23004	hypothetical prote
44	33	89.2	1308	2 A47253	epidermal growth f
45	33	89.2	1711	2 T21432	hypothetical prote

ALIGNMENTS

RESULT 1

T24298  
hypothetical protein T01E8.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T24298  
R:McMurray, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: Z19871  
A:Accession: T24298  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-367 <WIL>  
A:Cross-references: UNIPROT:Q22069; EMBL:Z48809; PIDN:CAA88744.1; GSPDB:GN00020; CESP:TO1E8  
C:Genetics:  
A:Gene: CESP:TO1E8.2  
A:Map position: 2  
A:Introns: 48/2; 200/2; 254/3

Query Match 100.0%; Score 37; DB 2; Length 367;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6  
|||||  
DB 6 TPPPAY 11

RESULT 2

S55019  
MAD polypeptide - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: S55019  
R:Seckelsky, J.J.; Newfeld, S.J.; Raftery, L.A.; Chartoff, E.H.; Gelbart, W.M.  
Genetics 139, 1347-1358, 1995  
A:Title: Genetic characterization and cloning of Mothers against dpp, a gene required for  
A:Reference number: S55018; MUID:95286061; PMID:7768443  
A:Accession: S55019  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-455 <SEK>  
A:Cross-references: UNIPROT:P42003; EMBL:U10328; NID:g551488; PIDN:AAB60230.1; PID:g551488  
C:Genetics:  
A:Gene: FlyBase:Mad  
A:Cross-references: FlyBase:FBgn0011648  
C:Superfamily: human transcription regulator MAD-4

Query Match 100.0%; Score 37; DB 2; Length 455;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6  
 Db 220 TPPPAY 225  
 RESULT 3  
 S68987  
 N:transcription activator Smad1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
 C:Accession: S68987; S71810  
 Nature 381, 620-623, 1996  
 A:Title: A human Mad protein acting as a BMP-regulated transcriptional activator.  
 A:Reference number: S68987; MUID:96238866; PMID:8637600  
 A:Accession: S68987  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-465 <LIU>  
 A:Cross-references: UNIPROT:Q15797; EMBL:U59423; NID:gl438076; PIDN:AAB06852.1; PID:gl438076  
 R:Zhang, Y.; Peng, X.H.; Wu, R.Y.; Derynck, R.  
 Nature 393, 168-172, 1996  
 A:Title: Receptor-associated Mad homologues synergize as effectors of the TGF-beta response  
 A:Reference number: S71797; MUID:96371046; PMID:8774881  
 A:Accession: S71810  
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-465 <ZHA>  
 C:Superfamily: human transcription regulator MAD-4  
 C:Keywords: transcription regulation  
 Query Match 100.0%; Score 37; DB 2; Length 465;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db 222 TPPPAY 227  
 RESULT 4  
 D84471  
 Probable membrane transporter [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: D84471  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84471  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-520 <STO>  
 A:Cross-references: UNIPROT:Q9SHZ3; GB:AE002093; NID:g4662639; PIDN:AAD26910.1; GSPDB:GN000020  
 C:Genetics:  
 A:Gene: At2g05760  
 A:Map position: 2  
 C:Superfamily: xanthine permease pbux  
 Query Match 100.0%; Score 37; DB 2; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db 306 TPPPAY 311

RESULT 5  
 AI0134  
 Probable phage protein YPO1096 [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AI0134  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Fildes, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AI0134  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-556 <KUP>  
 A:Cross-references: UNIPROT:Q8ZH17; GB:AL590842; PIDN:CAC89940.1; PID:gi15979164; GSPDB:GN000020  
 C:Genetics:  
 A:Gene: YPO1096  
 Query Match 100.0%; Score 37; DB 2; Length 556;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db 132 TPPPAY 137  
 RESULT 6  
 I51682  
 Epithelial sodium channel alpha subunit - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C:Accession: I51682  
 R:Puoti, A.; May, A.; Canessa, C.M.; Horiisberger, J.  
 Am. J. Physiol. 269, 188-197, 1995  
 A:Title: The highly selective, low conductance epithelial sodium channel of Xenopus laevis  
 A:Reference number: I51682  
 A:Accession: I51682  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-632 <PUO>  
 A:Cross-references: UNIPROT:P51167; EMBL:U23535; NID:g968935; PIDN:AAA74970.1; PID:g968935  
 C:Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I repeat F:355-390/Domain: fibronectin type I repeat homology <IF1>  
 C:Gene: alphaENaC  
 Query Match 100.0%; Score 37; DB 2; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPPPAY 6  
 Db 594 TPPPAY 599  
 RESULT 7  
 T20528  
 Hypothetical protein F07A11.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T20528  
 R:Palmer, S.  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z19287  
 A:Accession: T20528  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1095 <WIL>  
 A:Cross-references: UNIPROT:Q19132; EMBL:Z66511; PIDN:CAA91317.1; GSPDB:GN000020; CESP:F07A11  
 A:Experimental source: clone F07A11



## C:Genetics:

A:Gene: CBSP:F07A11.4

A:Map position: 2

A:Introns: 32/2; 202/2; 404/3; 587/3; 794/3; 964/1; 1004/3; 1080/2

## Query Match

Best Local Similarity 100.0%; Score 37; DB 2; Length 1095;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 316 TPPPAY 321

## RESULT 8

T41925

hypothetical protein U24 - human herpesvirus 7 (strain J1)

C:Species: human herpesvirus 7

A:Variety: strain J1

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T41925

R:Nicholas, J.

Submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of human H

A:Reference number: Z2022

A:Accession: T41925

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-82 &lt;NIC&gt;

A:Cross-references: UNIPROT:Q69505; EMBL:U43400; PIDN:AA054685.1

A:Experimental source: strain J1

C:Genetics:

A:Note: U24

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 82;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 5 TPPPSY 10

## RESULT 9

T44170

hypothetical protein U24 [imported] - human herpesvirus 6 (strain Z29)

C:Species: human herpesvirus 6

A:Variety: strain Z29

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Jun-2000

C:Accession: T44170

R:Dominguez, G.; Dambaugh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Fellett, P.E.

J. Virol. 73, 8040-8052, 1999

A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with human

A:Reference number: Z22734; MUID:99412318; PMID:10482553

A:Accession: T44170

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-88 &lt;DOM&gt;

A:Cross-references: EMBL:AF157706; PIDN:AA049637.1

A:Experimental source: strain Z29; variant B

C:Genetics:

A:Note: U24

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 88;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 6 TPPPSY 11

## RESULT 10

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 219;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

## T43984

hypothetical protein U24 [imported] - human herpesvirus 6 (strain HST)

C:Species: human herpesvirus 6

A:Variety: strain HST

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T43984

R:Isigawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawai

J. Virol. 73, 8053-8063, 1999

A:Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and

A:Reference number: Z22732; MUID:99412319; PMID:10482554

A:Accession: T43984

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-88 &lt;ISE&gt;

A:Cross-references: UNIPROT:Q9WT40; EMBL:AB021506; NID:g4995977; PIDN:BAA78245.1; PID:g49

A:Experimental source: strain HST; pop. variant B

C:Genetics:

A:Note: U24

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 88;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 6 TPPPSY 11

## RESULT 11

T09318

EoLF1 protein - human herpesvirus 6 (strain U1102)

C:Species: human herpesvirus 6

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T09318

R:Nicholas, J.; Martin, M.

J. Virol. 68, 597-610, 1994

A:Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of hu

A:Reference number: Z16644; MUID:94118404; PMID:8289364

A:Accession: T09318

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-93 &lt;NIC&gt;

A:Cross-references: UNIPROT:Q69048; EMBL:L25528; NID:g451932; PIDN:AAA16731.1; PID:g45194

C:Genetics:

A:Gene: EoLF1

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 93;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

Db 12 TPPPSY 17

## RESULT 12

BTEB2

protein - human

C:Species: Homo sapiens (man)

C:Date: 09-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C:Accession: S35643

R:Sogawa, K.; Imataka, H.; Yamasaki, Y.; Kusume, H.; Abe, H.; Fujii-Kuriyama, Y.

Nucleic Acids Res. 21, 1527-1532, 1993

A:Title: cDNA cloning and transcriptional properties of a novel GC box-binding protein, I

A:Reference number: S35643; MUID:93241930; PMID:8479902

A:Accession: S35643

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-219 &lt;SOG&gt;

A:Cross-references: UNIPROT:Q13887; GB:D14520; NID:g303596; PIDN:BAA03393.1; PID:d100390

## Query Match

Best Local Similarity 91.9%; Score 34; DB 2; Length 219;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

|||||

DB 85 TPPPSY 90

#### RESULT 13

hypothetical protein CPj0241 [imported] - Chlamydothila pneumoniae (strain J138)  
 A:Accession: A86521  
 C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: A86521  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; PMID:20330349; PMID:10871362  
 A:Accession: A86521  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-384 <STO>  
 A:Cross-references: UNIPROT:Q9JSH3; GB:BA000008; NID:G8978614; PIDN:BA098451.1; GSPDB:GN  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: CPj0241

Query Match 91.9%; Score 34; DB 2; Length 384;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

|||||

DB 335 TPPPSY 340

#### RESULT 14

hypothetical protein CP0521 [imported] - Chlamydothila pneumoniae (strains CWL029 and AR  
 C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: E72100; B81568  
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; PMID:99206606; PMID:10192388  
 A:Accession: E72100  
 A:Molecule type: DNA  
 A:Residues: 1-384 <ARN>  
 A:Cross-references: UNIPROT:Q9Z8U3; GB:AE001610; GB:AE001363; NID:g4376515; PIDN:AAD1839  
 A:Experimental source: strain CWL029  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
 A:Reference number: AB1500; PMID:20150255; PMID:10684935  
 A:Accession: B81568  
 A:Molecule type: DNA  
 A:Residues: 1-384 <REA>  
 A:Cross-references: GB:AE002212; GB:AE002161; NID:g7189430; PIDN:AAF38347.1; PID:g718943  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CPn0241; CP0521

Query Match 91.9%; Score 34; DB 2; Length 384;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

|||||

DB 335 TPPPSY 340

#### RESULT 15

A28913

regulatory protein brlA - Emericella nidulans

C:Species: Emericella nidulans, Aspergillus nidulans

C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004

C:Accession: A28913

R:Adams, T.H.; Boylan, M.T.; Timberlake, W.E.

Cell 54, 353-362, 1988

A:Title: brlA is necessary and sufficient to direct conidiophore development in Aspergill

A:Reference number: A28913; PMID:88282543; PMID:3293800

A:Accession: A28913

A:Molecule type: DNA

A:Residues: 1-432 <ADA>

A:Cross-references: UNIPROT:P10069

C:Genetics:

A:Gene: brlA

C:Keywords: DNA binding; zinc finger

Query Match 91.9%; Score 34; DB 2; Length 432;

Best Local Similarity 83.3%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6

|||||

DB 222 TPPPSY 227

Search completed: October 13, 2005, 15:11:25

Job time : 9.5055 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:03:23 ; Search time 11.4286 Seconds  
(without alignments)  
268.842 Million cell updates/sec

Title: US-09-385-918-16  
Perfect score: 37  
Sequence: 1 TPPPAY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	37	2 Q96AU8	Q96AU8 homo sapien
2	37	100.0	42	2 Q8AV30	Q8AV30 brachydanio
3	37	100.0	94	2 Q85B4	Q85B4 equus caball
4	37	100.0	116	2 Q8C6T0	Q8C6T0 mus musculus
5	37	100.0	170	2 Q9SM60	Q9SM60 ovis aries
6	37	100.0	203	2 Q8PKB9	Q8PKB9 xanthomonas
7	37	100.0	216	2 Q8IV31	Q8IV31 homo sapien
8	37	100.0	216	2 Q8NC22	Q8NC22 homo sapien
9	37	100.0	216	2 Q6ZME2	Q6ZME2 homo sapien
10	37	100.0	219	2 Q8MKC3	Q8MKC3 ovis aries
11	37	100.0	245	2 Q6IRI8	Q6IRI8 rattus norv
12	37	100.0	256	2 Q8ABH6	Q8ABH6 bacteroides
13	37	100.0	291	2 Q9I9L8	Q9I9L8 gallus gall
14	37	100.0	299	2 Q9W6N2	Q9W6N2 gallus gall
15	37	100.0	334	2 Q6NW84	Q6NW84 brachydanio
16	37	100.0	348	2 Q8JA33	Q8JA33 coxiella bu
17	37	100.0	367	2 Q22069	Q22069 caenorhabdi
18	37	100.0	377	2 Q9CRR7	Q9CRR7 mus musculus
19	37	100.0	393	2 Q7S2C7	Q7S2C7 neurospora
20	37	100.0	406	2 Q9W6N3	Q9W6N3 gallus gall
21	37	100.0	422	2 Q8T6S2	Q8T6S2 hydra atten
22	37	100.0	428	2 Q8C3Y6	Q8C3Y6 mus musculus
23	37	100.0	443	2 Q86NL5	Q86NL5 drosophila
24	37	100.0	455	1 MAD_DROME	P42003 drosophila
25	37	100.0	457	2 Q7Q8L8	Q7Q8L8 anopheles g
26	37	100.0	464	1 SMA5_BRARE	Q9W7e7 brachydanio
27	37	100.0	464	2 Q9I693	Q9I693 xenopus lae
28	37	100.0	464	2 Q9I913	Q9I913 xenopus lae
29	37	100.0	464	2 Q6DIS7	Q6DIS7 xenopus tro
30	37	100.0	464	2 Q6PF32	Q6PF32 xenopus lae
31	37	100.0	464	2 Q7ZU01	Q7ZU01 brachydanio

32 37 100.0 465 1 SMA1\_COTJA Q9I962 coturnix co  
33 37 100.0 465 1 SMA1\_HUMAN QI5797 homo sapien  
34 37 100.0 465 1 SMA1\_MOUSE P70340 mus musculus  
35 37 100.0 465 1 SMA5\_HUMAN Q99717 homo sapien  
36 37 100.0 465 1 SMA5\_MOUSE P97454 mus musculus  
37 37 100.0 465 1 SMA5\_RAT Q9RIV3 rattus norv  
38 37 100.0 465 2 Q6I9T1 Q6I9T1 homo sapien  
39 37 100.0 465 2 Q864V7 Q864V7 sus scrofa  
40 37 100.0 467 2 P79947 P79947 xenopus lae  
41 37 100.0 468 1 SMA1\_RAT P97588 rattus norv  
42 37 100.0 468 2 Q6P7A6 Q6P7A6 brachydanio  
43 37 100.0 472 1 SMA1\_BRARE Q9I8V2 brachydanio  
44 37 100.0 473 2 Q7T082 Q7T082 brachydanio  
45 37 100.0 474 2 Q6PUC3 Q6PUC3 anopheles g

## ALIGNMENTS

RESULT 1  
Q96AU8 PRELIMINARY; PRT; 37 AA.  
AC Q96AU8;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=Bone marrow;  
RC Strausberg R.;  
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC016719; AAH16719.1; -  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 37 AA; 4001 MW; A1EAA3B5B5E6734C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6  
Db 11 TPPPAY 16

RESULT 2  
Q8AV30 PRELIMINARY; PRT; 42 AA.  
AC Q8AV30;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Smad5 (Fragment).  
GN Name=smad5;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2263387; PubMed=12376102; DOI=10.1016/S0012-1606(02)90805-5;  
RA Kramer C., Mayr T., Nowak M., Bauer H., Schumacher J., Runke G.,  
RA Wagner D.S., Schmid B., Imai Y., Talbot W.S., Mullins M.C.,  
RA Hammerschmidt M.  
RT "Maternally supplied Smad5 is required for ventral specification in  
zebrafish embryos prior to zygotic Bmp signaling.";  
RL Dev. Biol. 250:263-279(2002).

```

DR EMBL; AY135144; AAN34935.1; -.
DR EMBL; AY135143; AAN34935.1; JOINED.
FT NON_TER 1
FT NON_TER 42
SQ SEQUENCE 42 AA; 4263 MW; BLE82C53C0A7A01E CRC64;

Query Match      100.0%; Score 37; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db |||||
27 TPPPAY 32

RESULT 3
Q865B4 PRELIMINARY; PRT; 94 AA.
AC Q865B4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sma45 (Fragment).
GN Name=SWAD5;
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tendon;
RA Arai K., Kasashima Y., Kuwano A., Yoshihara T.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB106114; BAC66060.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10078 MW; 94683F77EF069EA5 CRC64;

Query Match      100.0%; Score 37; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db |||||
48 TPPPAY 53

RESULT 4
Q8C6T0 PRELIMINARY; PRT; 116 AA.
AC Q8C6T0;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
DE library, clone:E030041A20 product:hypothetical protein, full insert
DE sequence.
GN Name=A930027H06Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

```

```

RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imokani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK053212; BAC35313.1; -.
DR MGD; MGI:1924444; A930027H06Rik.
KW Hypothetical protein.
SQ SEQUENCE 116 AA; 12590 MW; E2D2F6B2DA8ED17D CRC64;

Query Match      100.0%; Score 37; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
Db |||||
90 TPPPAY 95

RESULT 5
Q95M60 PRELIMINARY; PRT; 170 AA.
AC Q95M60;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sma41 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RN SEQUENCE FROM N.A.
RA Souza C.J.H., Nicol L., McNeilly A.S., Baird D.T.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY035385; AAK61393.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003619; Dwarfin_A.
DR Pfam; PF03165; MHL; 1.
DR SMART; SM00523; DWA; 1.
FT NON_TER 1
FT NON_TER 170
SQ SEQUENCE 170 AA; 19220 MW; 2C1D20CCDBEE7F3 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 160 TPPPAY 165

RESULT 6
QBPKB9
ID Q8PKB9 PRELIMINARY; PRT; 203 AA.
AC Q8PKB9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein XAC2256.
GN OrderedLocNames=XAC2256;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.N., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kiteajina J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463 (2002).
DR EMBL; AE011862; XAM37109.1; -.
KW Complete proteome.
SQ SEQUENCE 203 AA; 21411 MW; 68FD16EBE83E413B CRC64;

Query Match 100.0%; Score 37; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 36 TPPPAY 41

```

```

RESULT 7
Q8IV31
ID Q8IV31 PRELIMINARY; PRT; 216 AA.
AC Q8IV31
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein FLJ90586 (VPMH1932).
GN Name=FLJ90586; ORFNames=UNQ1932;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270 (2003).
DR EMBL; BC035517; AAK35517.1; -.
DR EMBL; AY358781; AAK9141.1; -.
KW Hypothetical protein.
SQ SEQUENCE 216 AA; 23729 MW; DF5A7DB1E4126063 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPAY 6
Db 190 TPPPAY 195

```

```

RESULT 8
Q8NC22 PRELIMINARY; PRT; 216 AA.
AC Q8NC22;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90586.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Matsuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075067; BAC11382.1; --
SQ SEQUENCE 216 AA; 23669 MW; 5D5A7DB1B4126076 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
DB 190 TPPPAY 195

RESULT 9
Q6ZME2 PRELIMINARY; PRT; 216 AA.
AC Q6ZME2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ23977.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK172816; BAD18786.1; --
SQ SEQUENCE 216 AA; 23671 MW; DF5A7DB1B411AF63 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
DB 190 TPPPAY 195

RESULT 10
Q8MKC3 PRELIMINARY; PRT; 219 AA.
AC Q8MKC3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Smad5 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza C.J.H., Baird D.T.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF508027; AAM34248.1; --
DR HSSP; Q15797; 1KHU.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00524; DWB; 1.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 24101 MW; BAB5739DE951B7C3 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6
DB 89 TPPPAY 94

RESULT 11
Q6IRI8 PRELIMINARY; PRT; 245 AA.
AC Q6IRI8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070905; AAH70905.1; --
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003619; Dwarfin_A.

```

```
DR Pfam; PF03165; MH1; 1.  
DR SMART; SM00523; DWA; 1.  
KW Hypothetical protein.  
FT NON_TER 245 245  
SQ SEQUENCE 245 AA; 27450 MW; 281008651BF043C4 CRC64;  
  
Query Match 100.0%; Score 37; DB 2; Length 245;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TPPPAY 6  
DB 221 TPPPAY 226  
  
RESULT 12  
Q8A8H6 PRELIMINARY; PRT; 256 AA.  
ID Q8A8H6  
AC Q8A8H6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames-BT1191;  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Bacteroidaceae; Bacteroides.  
OX NCBI_TaxID=818;  
RN [1]  
RP SEQUENCE FROM N.A. ATCC 29148;  
RC STRAIN=VPI-5482 / ATCC 29148;  
RX MEDLINE=22550958; PubMed=12663928; DOI=10.1126/science.1080029;  
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
RA Chiang H.C., Hooper L.V., Gordon J.I.;  
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";  
RL Science 299:2074-2076(2003).  
DR EMBL; AE016930; AAO76298.1; -.  
KW Complete proteome.  
SQ SEQUENCE 256 AA; 28486 MW; F4963B890947089D CRC64;  
  
Query Match 100.0%; Score 37; DB 2; Length 256;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TPPPAY 6  
DB 201 TPPPAY 206  
  
RESULT 13  
Q919L8 PRELIMINARY; PRT; 291 AA.  
ID Q919L8  
AC Q919L8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE BMP signal transducer Smad1 (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99456787; PubMed=10525349; DOI=10.1006/dbio.1999.9419;  
RA Yamada M., Szendro P.I., Prokcha A., Schwartz R.J., Eichele G.;  
RT "Evidence for a role of Smad6 in chick cardiac development.";  
RL Dev. Biol. 215:48-61(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Szendro P.I., Grunenberg U., Eichele G.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF233238; AAF36983.1; -.
```

```
DR HSP; Q15797; 1KHU.  
DR TRANSFAC; T03889; -.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001132; Dwarf1n.  
DR InterPro; IPR003619; Dwarf1n.A.  
DR InterPro; IPR008984; SMAD_FHA.  
DR Pfam; PF03165; MH1; 1.  
DR Pfam; PF03166; MH2; 1.  
DR SMART; SM00523; DWA; 1.  
FT NON_TER 291 291  
SQ SEQUENCE 291 AA; 32755 MW; 7F9C39B17748E446 CRC64;  
  
Query Match 100.0%; Score 37; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TPPPAY 6  
DB 207 TPPPAY 212  
  
RESULT 14  
Q9W6N2 PRELIMINARY; PRT; 299 AA.  
ID Q9W6N2  
AC Q9W6N2;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Smad1 protein (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart;  
RA Lough J.W., Vincent E.B., Weeks D.L.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF143239; AAD30150.1; -.  
DR HSP; Q15797; 1KHU.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001132; Dwarf1n.  
DR InterPro; IPR008984; SMAD_FHA.  
DR Pfam; PF03165; MH2; 1.  
DR SMART; SM00524; DWA; 1.  
FT NON_TER 1  
SQ SEQUENCE 299 AA; 33210 MW; 28232FE701D8233B CRC64;  
  
Query Match 100.0%; Score 37; DB 2; Length 299;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TPPPAY 6  
DB 56 TPPPAY 61  
  
RESULT 15  
Q6NW84 PRELIMINARY; PRT; 334 AA.  
ID Q6NW84  
AC Q6NW84;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.
```

OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Strausberg R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; BC067685; AAH67685.1; -.  
DR HSSP; P10037; 1AU7.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain like.  
DR InterPro; IPR010982; Lambda\_like\_DNA.  
DR InterPro; IPR000327; POU.  
DR InterPro; IPR007103; POU\_homeo.  
DR Pfam; PF00046; Homeobox; 1.  
DR Pfam; PF00157; Pou; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRINTS; PR00028; POU\_DOMAIN.  
DR ProDom; PD000010; Homeobox; 1.  
DR ProDom; PD000583; POU; 1.  
DR SMART; SM00389; HOX; 1.  
DR SMART; SM00352; POU; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
DR PROSITE; PS00035; POU\_1; 1.  
DR PROSITE; PS00463; POU\_2; 1.  
KW DNA-binding; Homeobox; Hypothetical protein; Nuclear protein.  
FT NON TER 1  
SQ SEQUENCE 334 AA; 36699 MW; 25031A1FD4FEBD41 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPAY 6  
Db 21 TPPPAY 26  
|||||

Search completed: October 13, 2005, 15:09:28  
Job time : 13.4286 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:01:04 ; Search time 13.5165 Seconds  
(without alignments)  
171.684 Million cell updates/sec

Title: US-09-385-918-18

Perfect score: 39

Sequence: 1 TPPPGY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	6	AAB83033	Human Sma
2	39	100.0	11	AAB83034	Human Sma
3	39	100.0	14	AAB83039	Human Sma
4	39	100.0	14	AAB83040	Human Sma
5	39	100.0	16	AAB82838	Human Sma
6	39	100.0	16	AAB82837	Human Sma
7	39	100.0	381	AAB82837	Human Sma
8	39	100.0	417	AAB80934	Human Sma
9	39	100.0	424	AAB80934	Human Sma
10	39	100.0	424	AAB80936	Human Sma
11	39	100.0	424	ABP53783	Human Sma
12	39	100.0	425	ABP53783	Human Sma
13	39	100.0	425	AAW55965	Human Sma
14	39	100.0	425	AAW55965	Human Sma
15	39	100.0	425	AAW55965	Human Sma
16	39	100.0	425	AAW55965	Human Sma
17	39	100.0	425	ABP53784	Human Sma
18	39	100.0	425	ABP53784	Human Sma
19	39	100.0	425	ADP34562	Human Sma
20	39	100.0	425	ADP34562	Human Sma
21	39	100.0	425	ADP34562	Human Sma
22	39	100.0	425	ADP34562	Human Sma
23	39	100.0	425	ADP34562	Human Sma
24	39	100.0	425	ADP34562	Human Sma
25	39	100.0	425	ADP34562	Human Sma

26	39	100.0	467	2	AAW56692	Mouse tum
27	39	100.0	467	2	AAW56691	Human tum
28	39	100.0	467	2	AAW90239	Xenopus 1
29	39	100.0	467	2	AAW90241	Human Sma
30	39	100.0	467	2	AAW82611	Murine Sma
31	39	100.0	467	2	AAW82613	Murine Sma
32	39	100.0	467	2	AAW82610	Murine Sma
33	39	100.0	467	2	AAW82609	Murine Sma
34	39	100.0	467	2	AAW82612	Murine Sma
35	39	100.0	467	2	AAW82595	Mouse Sma
36	39	100.0	467	2	AAW82608	Murine Sma
37	39	100.0	467	2	AAW82614	Murine Sma
38	39	100.0	467	2	AAW82607	Human Sma
39	39	100.0	467	2	AAW82607	Human Sma
40	39	100.0	467	5	ABP53781	Xenopus S
41	39	100.0	467	5	ABP53782	Human Sma
42	39	100.0	467	7	ADP54507	Rat Prote
43	39	100.0	467	7	ADP54505	Human Pro
44	39	100.0	467	7	ADP54501	Human Pro
45	39	100.0	467	7	ADP54503	Rat Prote

## ALIGNMENTS

RESULT 1

AAB83033

ID AAB83033 standard; peptide; 6 AA.

XX AC AAB83033;

XX 25-JUN-2001 (first entry)

XX Human Smad 2 and Smad 3 PY motif consensus sequence.

XX Human: HECT; homologous to E6 carboxyl terminus; ubiquitination;

XX HECT E3 ubiquitin ligase; antiinflammatory; cytosolic; osteopontin;

XX antimicrobial; neuroprotective; transforming growth factor beta;

XX TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;

XX inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;

XX renal failure; neurodegeneration; fibrosis; PY motif; Smad 2; Smad 3.

XX Homo sapiens.

XX WO200116604-A1.

XX 08-MAR-2001.

XX 29-AUG-2000; 2000WO-US023729.

XX 30-AUG-1999; 99US-00385918.

XX (SIGN-) SIGNAL PHARM INC.

XX Hoekstra MF, Xie W, Murray BW, Mercurio FM;

XX WPI; 2001-327913/34.

XX Screening for modulators of TGF-beta and/or bone morphogenic protein

XX (BMP) mediated signaling useful for treating cancer and osteoporosis by

XX evaluating the ability of agents to modulate Smad protein degradation.

XX Claim 4; Page 35; 75pp; English.

XX The present sequence is the Smad PY motif consensus sequence. The PY

XX motif binds to the WW domain of HECT (homologous to E6 carboxyl terminus)

XX E3 ubiquitin ligase, resulting in ubiquitination of Smad by the E3

XX ubiquitin ligase. The sequence is provided in a specification relating to

XX a method for screening for agents that modulate transforming growth

XX factor (TGF)-beta and/or bone morphogenic protein (BMP)-mediated

XX signalling. The method involves evaluating the effect of an agent on

XX binding of HECT E3 ubiquitin ligase WW domain to Smad PY motif, on

XX ubiquitination of Smad protein by E3 ubiquitin ligase, or on the cellular

CC levels of Smad protein HECT E3 ubiquitin ligase activity. The method is  
 CC useful for stimulating bone formation in a patient or treating a  
 CC condition associated with insufficient TGF-beta and/or BMP-mediated cell  
 CC signalling. Agents that inhibit BMP-mediated signalling are useful for  
 CC treating inflammation, ageing, cancer and infectious diseases. Agents  
 CC that augment BMP-mediated signalling are useful for stimulating bone  
 CC anabolism as well as treating broken bones, osteoporosis, and acute or  
 CC chronic renal failure. Agents that inhibit TGF-mediated signalling are  
 CC useful for treating cancer, inflammation, neurodegeneration and fibrosis  
 CC  
 XX Sequence 6 AA;

Query Match 100.0%; Score 39; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 |||||  
 Db 1 TTPPGY 6

RESULT 2  
 AAB83034  
 ID AAB83034 standard; peptide; 11 AA.

XX AC AAB83034;

DT 25-JUN-2001 (first entry)

XX DE Human Smad 2 and Smad 3 PY motif consensus sequence.

XX KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
 KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
 KW antimicrobial; neuroprotective; transforming growth factor beta;  
 KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
 KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
 KW renal failure; neurodegeneration; fibrosis; PY motif; Smad 2; Smad 3.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 7  
 FT /label= Ile, Leu

XX PN WO200116604-A1.

XX PD 08-MAR-2001.

XX PF 29-AUG-2000; 2000WO-US023729.

XX PR 30-AUG-1999; 99US-00385918.

XX PA (SIGN-) SIGNAL PHARM INC.

XX PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;

XX DR WPI; 2001-327913/34.

XX PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
 PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
 PT evaluating the ability of agents to modulate Smad protein degradation.

XX PS Disclosure; Page 14; 75pp; English.

XX CC The present sequence is a Smad PY motif consensus sequence. The PY motif  
 CC binds to the WW domain of HECT (homologous to E6 carboxyl terminus) E3  
 CC ubiquitin ligase, resulting in ubiquitination of Smad by the E3 ubiquitin  
 CC ligase. The sequence is provided in a specification relating to a method  
 CC for screening for agents that modulate transforming growth factor (TGF)-  
 CC beta and/or bone morphogenic protein (BMP)-mediated signalling. The  
 CC method involves evaluating the effect of an agent on binding of HECT E3  
 CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
 CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein

CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
 CC bone formation in a patient or treating a condition associated with  
 CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
 CC inhibit BMP-mediated signalling are useful for treating inflammation,  
 CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
 CC signalling are useful for stimulating bone anabolism as well as treating  
 CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
 CC that inhibit TGF-mediated signalling are useful for treating cancer,  
 CC inflammation, neurodegeneration and fibrosis  
 XX Sequence 11 AA;

Query Match 100.0%; Score 39; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 |||||  
 Db 1 TTPPGY 6

RESULT 3  
 AAB83039  
 ID AAB83039 standard; peptide; 14 AA.

XX AC AAB83039;

DT 25-JUN-2001 (first entry)

XX DE Human Smad PY motif #5.

XX KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
 KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
 KW antimicrobial; neuroprotective; transforming growth factor beta;  
 KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
 KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
 KW renal failure; neurodegeneration; fibrosis; PY motif.

XX OS Homo sapiens.

XX PN WO200116604-A1.

XX PD 08-MAR-2001.

XX PF 29-AUG-2000; 2000WO-US023729.

XX PR 30-AUG-1999; 99US-00385918.

XX PA (SIGN-) SIGNAL PHARM INC.

XX PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;

XX DR WPI; 2001-327913/34.

XX PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
 PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
 PT evaluating the ability of agents to modulate Smad protein degradation.

XX PS Example 2; Page 31; 75pp; English.

XX CC The present sequence is a Smad PY motif. The PY motif binds to the WW  
 CC domain of HECT (homologous to E6 carboxyl terminus) E3 ubiquitin ligase,  
 CC resulting in ubiquitination of Smad by the E3 ubiquitin ligase. The  
 CC sequence is provided in a specification relating to a method for  
 CC screening for agents that modulate transforming growth factor (TGF)-beta  
 CC and/or bone morphogenic protein (BMP)-mediated signalling. The method  
 CC involves evaluating the effect of an agent on binding of HECT E3  
 CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
 CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein  
 CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
 CC bone formation in a patient or treating a condition associated with  
 CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
 CC inhibit BMP-mediated signalling are useful for treating inflammation,

CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
 CC signalling are useful for stimulating bone anabolism as well as treating  
 CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
 CC that inhibit TGF-mediated signalling are useful for treating cancer,  
 CC inflammation, neurodegeneration and fibrosis  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 39; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
 Db 4 TPPPGY 9

RESULT 4  
 AAB83040  
 ID AAB83040 standard; peptide; 14 AA.  
 XX AC  
 XX AAB83040;  
 XX  
 DT 25-JUN-2001 (first entry)  
 XX  
 DE Human Smad PY motif #6.  
 XX  
 KW Human; HECT; homologous to E6 carboxyl terminus; ubiquitination;  
 KW HECT E3 ubiquitin ligase; antiinflammatory; cytostatic; osteopathic;  
 KW antimicrobial; neuroprotective; transforming growth factor beta;  
 KW TGF-beta; bone morphogenic protein; BMP; Smad; cell signalling;  
 KW inflammation; ageing; cancer; infection; bone anabolism; osteoporosis;  
 KW renal failure; neurodegeneration; fibrosis; PY motif.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200116604-A1.  
 XX  
 PD 08-MAR-2001.  
 XX  
 PF 29-AUG-2000; 2000WO-US023729.  
 XX  
 PR 30-AUG-1999; 99US-00385918.  
 XX  
 PA (SIGN-) SIGNAL PHARM INC.  
 XX  
 PI Hoekstra MF, Xie W, Murray BW, Mercurio FM;  
 XX  
 DR WPI; 2001-327913/34.  
 XX  
 PT Screening for modulators of TGF-beta and/or bone morphogenic protein  
 PT (BMP) mediated signaling useful for treating cancer and osteoporosis by  
 PT evaluating the ability of agents to modulate Smad protein degradation.  
 XX  
 PS Example 2; Page 31; 75pp; English.  
 XX  
 CC The present sequence is a Smad PY motif. The PY motif binds to the WW  
 CC domain of HECT (homologous to E6 carboxyl terminus) E3 ubiquitin ligase,  
 CC resulting in ubiquitination of Smad by the E3 ubiquitin ligase. The  
 CC sequence is provided in a specification relating to a method for  
 CC screening for agents that modulate transforming growth factor (TGF)-beta  
 CC and/or bone morphogenic protein (BMP)-mediated signalling. The method  
 CC involves evaluating the effect of an agent on binding of HECT E3  
 CC ubiquitin ligase WW domain to Smad PY motif, on ubiquitination of Smad  
 CC protein by E3 ubiquitin ligase, or on the cellular levels of Smad protein  
 CC HECT E3 ubiquitin ligase activity. The method is useful for stimulating  
 CC bone formation in a patient or treating a condition associated with  
 CC insufficient TGF-beta and/or BMP-mediated cell signalling. Agents that  
 CC inhibit BMP-mediated signalling are useful for treating inflammation,  
 CC ageing, cancer and infectious diseases. Agents that augment BMP-mediated  
 CC signalling are useful for stimulating bone anabolism as well as treating  
 CC broken bones, osteoporosis, and acute or chronic renal failure. Agents  
 CC that inhibit TGF-mediated signalling are useful for treating cancer,

CC inflammation, neurodegeneration and fibrosis  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 39; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
 Db 4 TPPPGY 9

RESULT 5  
 AAW82838  
 ID AAW82838 standard; peptide; 16 AA.  
 XX AC  
 XX AAW82838;  
 XX  
 DT 01-FEB-1999 (first entry)  
 XX  
 DE PY motif of the human MAD-3 protein.  
 XX  
 KW Endothelial MAD interactor protein 1; mothers against dpp; MAD-3;  
 KW decapentaplegic; human; transforming growth factor-beta; TGF-beta;  
 KW EM11 deficiency; cardiovascular disease; atherosclerosis; ischaemia;  
 KW reperfusion; hypertension; restenosis; arterial inflammation;  
 KW proliferative disease; cancer; colorectal; pancreatic; antisense;  
 KW angiogenesis; autoimmune disease; fibrosis; wound healing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9845467-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US007356.  
 XX  
 PR 10-APR-1997; 97US-00844312.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Gimeno CJ, Falb DA;  
 XX  
 DR WPI; 1998-583204/49.  
 XX  
 PT Nucleic acid encoding endothelial MAD interactor I - for modulating cell  
 PT proliferation and differentiation, e.g. in cases of atherosclerosis and  
 PT cancer of colon or pancreas.  
 XX  
 PS Example 4; Page 67; 95pp; English.  
 XX  
 CC The present sequence represents the PY motif of a MAD-3 (mothers against  
 CC dpp (decapentaplegic)) protein. The specification describes a human  
 CC endothelial MAD interactor 1 protein EM11. The protein modulates  
 CC transforming growth factor-beta (TGF-beta) response in TGF-beta  
 CC responsive cells. The EM11 protein is used to treat EM11 deficiency. They  
 CC are especially used for treatment of cardiovascular disease (specifically  
 CC atherosclerosis but also ischaemia/reperfusion, hypertension, restenosis  
 CC and arterial inflammation) or proliferative diseases, especially cancer  
 CC of gut-derived cells, e.g. colorectal or pancreatic cancer, also to  
 CC modulate angiogenesis, treat autoimmune disease or fibrosis and to  
 CC regulate wound healing  
 XX  
 SQ Sequence 16 AA;  
 Query Match 100.0%; Score 39; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
 Db 7 TPPPGY 12

```

RESULT 6
AAW82837
ID AAW82837 standard; peptide; 16 AA.
XX
AC AAW82837;
XX
DT 01-FEB-1999 (first entry)
XX
DE PY motif of the human MAD-2 protein.
XX
KW Endothelial MAD interactor protein 1; mothers against ddp; MAD-2;
KW decapentaplegic; human; transforming growth factor-beta; TGF-beta;
KW EM1 deficiency; cardiovascular disease; atherosclerosis; ischaemia;
KW reperfusion; hypertension; restenosis; arterial inflammation;
KW proliferative disease; cancer; colorectal; pancreatic; antisense;
KW angiogenesis; autoimmune disease; fibrosis; wound healing.
XX
OS Homo sapiens.
XX
PN WO9845467-A1.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US007356.
XX
PR 10-APR-1997; 97US-00844312.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gimeno CJ, Falb DA;
XX
DR WPI; 1998-583204/49.
XX
PT Nucleic acid encoding endothelial MAD interactor I - for modulating cell
PT proliferation and differentiation, e.g. in cases of atherosclerosis and
PT cancer of colon or pancreas.
XX
PS Example 4; Page 67; 95pp; English.
XX
CC The present sequence represents the PY motif of a MAD-2 (mothers against
CC ddp (decapentaplegic)) protein. The specification describes a human
CC endothelial MAD interactor 1 protein EM1. The protein modulates
CC transforming growth factor-beta (TGF-beta) response in TGF beta
CC responsive cells. The EM1 protein is used to treat EM1 deficiency. They
CC are especially used for treatment of cardiovascular disease (specifically
CC atherosclerosis but also ischaemia/reperfusion, hypertension, restenosis
CC and arterial inflammation) or proliferative diseases, especially cancer
CC of gut-derived cells, e.g. colorectal or pancreatic cancer, also to
CC modulate angiogenesis, treat autoimmune disease or fibrosis and to
CC regulate wound healing
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 39; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 7 TTPPGY 12

RESULT 7
AAW99363
ID AAW99363 standard; protein; 381 AA.
XX
AC AAW99363;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human SMAD3 splice variant protein.

```

```

XX Mothers against ddp; MAD; splice variant; signal transduction;
KW decapentaplegic; diagnosis; renal failure; atherosclerosis; fibrosis.
XX
OS Homo sapiens.
XX
PN EP894856-A1.
XX
PD 03-FEB-1999.
XX
PF 22-JUL-1998; 98EP-00305858.
XX
PR 01-AUG-1997; 97US-00904874.
XX
PA (SMIK ) SMITHKLINE BECKMAN CORP.
XX
PI Laping N, Zhu Y;
XX
DR WPI; 1999-108348/10.
DR N-PSDB; AAX25710.
XX
XX New Mothers against ddp (MAD) polypeptide (SMAD3d) and polynucleotide -
PT useful as diagnostic reagents and for prevention and treatment of chronic
PT renal failure, atherosclerosis and fibrosis.
XX
PS Claim 11; Page 7; 20pp; English.
XX
CC This sequence represents a novel Mothers against ddp (MAD) protein
CC encoded by a splice variant, which is required for the signal
CC transduction of DPP (decapentaplegic). The SMAD3d polypeptides and
CC polynucleotides are useful for diagnosing, preventing or treating
CC diseases including chronic renal failure, atherosclerosis and fibrosis
XX
SQ Sequence 381 AA;

Query Match 100.0%; Score 39; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 135 TTPPGY 140

RESULT 8
AAB80934
ID AAB80934 standard; protein; 417 AA.
XX
AC AAB80934;
XX
DT 04-JUN-2001 (first entry)
XX
DE Human Smad2 protein.
XX
KW Smad; Cytostatic; vulnery; cerebroprotective; immunosuppressive;
KW Anti-HIV; antidiabetic; ophthalmological; antiinflammatory; SIM;
KW Smad Interaction Motif; tissue repair; fibrotic condition; human;
KW immunosuppression; diabetic nephropathy; tumour.
XX
OS Homo sapiens.
XX
PN WO200114413-A2.
XX
PD 01-MAR-2001.
XX
PF 25-AUG-2000; 2000WO-GB003265.
XX
PR 25-AUG-1999; 99GB-00020000.
XX
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Germain SE, Hill CS, Howell MT;
XX

```

DR	WPI; 2001-265836/27.
XX	
PT	Polypeptide capable of interacting with a Smad polypeptide, useful in the
PT	treatment of cancer and for tissue remodelling or healing of a wound,
PT	injury or surgery, comprises a Smad Interaction Motif and is less than 32
PT	amino acids in length.
XX	
XX	Claim 6; Fig 12; 179pp; English.
PS	
PS	The present invention relates to peptides capable of interacting with a
CC	Smad protein, comprising a Smad Interaction Motif (SIM); amino acid
CC	sequence PP(T/N/K). The present sequence is a Smad protein. Smad proteins
CC	are a family of highly conserved, intracellular proteins that signal
CC	cellular responses downstream of Transforming Growth Factor-beta (TGF-
CC	beta) family serine/threonine kinase receptors. The SIM is thought to be
CC	necessary for interaction with the MH2 domain of Smad2. The Smad-
CC	interacting peptides of the present invention are useful in the
CC	manufacture of a medicament for the treatment of a patient in need of
CC	modulation of activin or TGF-beta signalling; cancer; a patient in need
CC	of reducing extracellular matrix deposition, encouraging tissue repair
CC	and/or regeneration, tissue remodelling or healing of a wound, injury or
CC	surgery, or reducing scar tissue formation arising from injury to the
CC	brain; a patient with or at risk of end-stage organ failure, pathologic
CC	extracellular matrix accumulation, a fibrotic condition, disease states
CC	associated with immunosuppression (such as different forms of malignancy,
CC	chronic degenerative diseases, and AIDS), diabetic nephropathy, tumour
CC	growth, kidney damage (for e.g. obstructive neuropathy, IGA nephropathy
CC	or non-inflammatory renal disease) or renal fibrosis
XX	
XX	
XX	Sequence 417 AA;
SQ	
	Query Match 100.0%; Score 39; DB 4; Length 417;
	Best Local Similarity 100.0%; Pred. No. 5.8e+02;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TTPPGY 6
DB	170 TTPPGY 175
RESULT 9	
AAW90243	
ID	AAW90243 standard; protein; 424 AA.
XX	
XX	AAW90243;
XX	
DT	10-MAY-1999 (first entry)
DE	
DE	Human Smad3 protein.
XX	
KW	Transforming growth factor-beta superfamily signalling; modulator; Smad2;
KW	TGF-beta; detection; FAST-1; MH2 domain; Smad interaction domain; SID;
KW	treatment; developmental; disorder; immunological; cancer; diagnosis;
KW	Smad3.
XX	
OS	Homo sapiens.
XX	
PN	WO9853830-A1.
XX	
XX	03-DEC-1998.
PD	
XX	28-MAY-1998; 98WO-US010983.
XX	
XX	28-MAY-1997; 97US-0047991P.
PR	
XX	(HARD ) HARVARD COLLEGE.
PA	
XX	Whitman M, Chen X;
PI	
XX	
XX	WPI; 1999-059773/05.
DR	
DR	N-PSDB; AAV72113.
XX	
XX	Modulating TGF-beta superfamily signalling - comprises use of compounds

CC beta) family serine/threonine kinase receptors. The SIM is thought to be  
 CC necessary for interaction with the MH2 domain of Smad2. The Smad-  
 CC interacting peptides of the present invention are useful in the  
 CC manufacture of a medicament for the treatment of a patient in need of  
 CC modulation of activin or TGF-beta signalling; cancer; a patient in need  
 CC of reducing extracellular matrix deposition, encouraging tissue repair  
 CC and/or regeneration, tissue remodelling or healing of a wound, injury or  
 CC surgery, or reducing scar tissue formation arising from injury to the  
 CC brain; a patient with or at risk of end-stage organ failure, pathologic  
 CC extracellular matrix accumulation, a fibrotic condition, disease states  
 CC associated with immunosuppression (such as different forms of malignancy,  
 CC chronic degenerative diseases, and AIDS), diabetic nephropathy, tumour  
 CC growth, kidney damage (for e.g. obstructive neuropathy, IGA nephropathy  
 CC or non-inflammatory renal disease) or renal fibrosis  
 XX  
 SQ Sequence 424 AA;  
 Query Match 100.0%; Score 39; DB 4; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TPPPGY 6  
 Db 178 TPPPGY 183  
 RESULT 11  
 ABP53783  
 ID ABP53783 standard; protein; 424 AA.  
 XX  
 AC ABP53783;  
 XX  
 DT 06-JAN-2003 (first entry)  
 XX  
 DE Human Smad3 (HSmad3) protein sequence.  
 XX  
 KW Smad; Smad interacting protein; XFast-3; Antennapedia helix 3; FAST;  
 KW activin response factor; ARF; cytotstatic; vulnerary; antidiabetic;  
 KW anti-HIV; nephrotropic; activin; transforming growth factor beta;  
 KW TGF-beta; signalling; cancer; tissue repair; immunosuppression;  
 KW malignancy; chronic degenerative disease; AIDS; diabetic nephropathy;  
 KW tumour growth; kidney damage; obstructive nephropathy;  
 KW immunoglobulin A nephropathy; IGA nephropathy; renal fibrosis;  
 KW non-inflammatory renal disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200270702-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 28-FEB-2002; 2002WO-GB000877.  
 XX  
 PR 01-MAR-2001; 2001GB-00005063.  
 XX  
 PA (INCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX  
 PI Hill C, Randall R, Howell M;  
 XX  
 DR WPI; 2002-732794/79.  
 XX  
 PT New interacting polypeptide capable of interacting with an activin  
 PT response factor component, useful in medicine or for the manufacture of a  
 PT medicament for treating diabetic nephropathy, tumor growth, kidney damage  
 PT or wound healing.  
 XX  
 PS Disclosure; Fig 3; 153pp; English.  
 XX  
 CC The present invention describes a polypeptide (I) capable of interacting  
 CC (i.e. interacting polypeptide) with an activin response factor (ARF)  
 CC component. The interacting polypeptide comprises: (a) a Fast motif 1  
 CC (FM1) and/or a Fast motif 2 (FM2), and is less than 84, 83, 82, 81, 80,  
 CC 75, 70, 65, 60, 55, 50, 45, 40, 35, 32, 31, or 30 amino acids in length;

CC (b) the amino acid sequence LPTSY (S1) and/or PNXXXP(P/L) (S2) or  
 CC PN(VA)VA(W)P(P/L) (S3) and is less than 84, 83, 82, 81, 80, 75, 70,  
 CC 65, 60, 55, 50, 45, 40, 35, 32, 31, or 30 amino acids in length; or (c) a  
 CC FM1 and/or FM2 motif, e.g. the sequence of S1, or 3 or 4 out of 5 of its  
 CC residues, and/or the sequence of S2 or S3, or 4, 5 or 6 out of 7 of its  
 CC residues, or out of 5 of its residues, and is not full-length Xenopus,  
 CC Zebrafish or human FAST1 or its fragment, mouse FAST2 or Xenopus FAST3.  
 CC (I) has cytostatic, vulnerary, anti-HIV, antidiabetic and nephrotropic  
 CC activities. (I) can have a transforming growth factor (TGF) agonist beta,  
 CC TGF antagonist beta, activin agonist or activin antagonist mechanism of  
 CC action. (I) can be used in a medicine or in the manufacture of a  
 CC medicament for treating a patient: (i) in need of modulation of activin,  
 CC TGF-beta or other superfamily member signalling; (ii) with cancer; (iii)  
 CC in need of reducing extracellular matrix deposition, encouraging tissue  
 CC repair and/or regeneration, tissue remodeling or healing of a wound,  
 CC injury or surgery, or reducing scar tissue formation arising from injury  
 CC to the brain; or (iv) with or at risk of end-stage organ failure,  
 CC pathological extracellular matrix accumulation, a fibrotic condition,  
 CC disease states associated with immunosuppression (e.g. different forms of  
 CC malignancy, chronic degenerative diseases or AIDS), diabetic nephropathy,  
 CC tumour growth, kidney damage (e.g. obstructive nephropathy,  
 CC immunoglobulin A (IGA) nephropathy or non-inflammatory renal disease) or  
 CC renal fibrosis. The present sequence represents a human Smad3 (HSmad3)  
 XX protein given in the exemplification of the present invention  
 XX  
 SQ Sequence 424 AA;  
 Query Match 100.0%; Score 39; DB 5; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TPPPGY 6  
 Db 178 TPPPGY 183  
 RESULT 12  
 AAW55965  
 ID AAW55965 standard; protein; 425 AA.  
 XX  
 AC AAW55965;  
 XX  
 DT 17-AUG-1998 (first entry)  
 XX  
 DE Amino acid sequence of human MADr3 protein.  
 XX  
 KW Human; MAD; Mothers against DPP; Drosophila decapentaplegic; MADr3;  
 KW transforming growth factor-beta; TGF-beta; signal transduction; BMP2;  
 KW BMP4; MADr4; wound healing; AIDS dementia; treatment; ocular disorder;  
 KW kidney; liver; prevention; scarring; ulcer; corneal incision; cancer;  
 KW renal failure; arthritis; osteoporosis; atherosclerosis;  
 KW polycystic kidney disease; congestive heart failure.  
 XX  
 OS Homo sapiens.  
 XX  
 FN EP837073-A1.  
 XX  
 PD 22-APR-1998.  
 XX  
 PF 15-OCT-1997; 97EP-00308160.  
 XX  
 PR 16-OCT-1996; 96US-00732028.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Laping NJ;  
 XX  
 DR WPI; 1998-219071/20.  
 DR N-PSDB; AAV26065.  
 XX  
 FT DNA encoding human MAD protein - useful for producing recombinant  
 FT protein, gene therapy, etc.  
 XX

PS Claim 5; Fig 1; 47pp; English.

XX The present sequence represents a human MAD (Mothers against DPP

CC (Drosophila decapentaplegic) (sic)) homologue, designated MADr3.

CC Different transforming growth factor-beta (TGF-beta) family members may

CC signal through different MAD isoforms. MAD proteins have been found to be

CC required for signal transduction of DPP and BMP2/BMP4. MADr3 and MADr4

CC polypeptides can be used to stimulate wound healing, to restore normal

CC neurological function after trauma or AIDS dementia, to treat ocular

CC disorders, to target certain cells, to treat kidney and liver disorders,

CC to prevent scarring, to treat ulcers and corneal incisions or to treat

CC cancer. MADr3 antagonists can be used to treat renal failure, wound

CC healing and prevention of scar formation, arthritis, osteoporosis,

CC atherosclerosis, polycystic kidney disease and congestive heart failure.

CC The polypeptides can be used to detect and diagnose disease

XX

SQ Sequence 425 AA;

Query Match 100.0%; Score 39; DB 2; Length 425;

Best Local Similarity 100.0%; Pred. No. 5.9e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPPPGY 6

Db 179 TPPPGY 184

|||||

RESULT 13

AAAY69623

ID AAY69623 standard; protein; 425 AA.

XX

AC AAY69623;

XX

DT 19-APR-2000 (first entry)

XX

DE Human Smad3.

XX

XX Smad3; MADH3; hMAD3; JV15-2; TGF-beta signalling pathway;

KW transcription factor; expression inhibition; antisense therapy;

KW tumour formation; inflammation.

XX

OS Homo sapiens.

XX

FN US6013788-A.

XX

PD 11-JAN-2000.

XX

PF 09-APR-1999; 99US-00289376.

XX

PR 09-APR-1999; 99US-00289376.

XX

PA (ISIS-) ISIS PHARM INC.

XX

PI Monia BP, Cowsett LM;

XX

DR WPI; 2000-126072/11.

XX

DR N-PSDB; AAZ59789.

XX

PT Antisense inhibition of the human Smad3 gene, useful for diagnosing,

PT preventing and treating conditions associated with Smad3 expression e.g.

PT inflammation.

XX

PS Disclosure; Col 41-44; 31pp; English.

XX

CC This sequence represents human Smad3. The Smad proteins are a family of

CC cytosolic proteins which are involved in TGF-beta superfamily signal

CC transduction. On ligand binding, TGF-beta superfamily proteins (such as

CC bone morphogenetic protein (BMP), activin and TGF-betas themselves)

CC phosphorylate Smad proteins, which then homo- or heterodimerise and

CC translocate to the nucleus to activate target gene transcription. Smad3

CC (also known as MADH3, hMAD3 and JV15-2) is a member of a subgroup of Smad

CC family transcription factors, the pathway- restricted Smads, which are

CC regulated by TGF-beta and activins. It can heterodimerise with Smad4

CC

CC (US6013787-A, AAY69622), the complex being able to activate TGF-beta

CC inducible transcription. The invention relates to antisense

CC oligonucleotides targeted to the human Smad3 gene, which inhibit its

CC expression. A series of oligonucleotides (AAZ59796-259735) were designed

CC to target different regions of the human Smad3 RNA, and were analysed for

CC their effect on Smad3 mRNA levels by quantitative real-time PCR. The

CC oligonucleotides of the invention are useful for diagnosis, prevention

CC and treatment of conditions associated with Smad3 expression, such as

XX tumour formation, inflammation and certain infections

SQ Sequence 425 AA;

Query Match 100.0%; Score 39; DB 3; Length 425;

Best Local Similarity 100.0%; Pred. No. 5.9e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPPPGY 6

Db 179 TPPPGY 184

|||||

RESULT 14

AAAY6726

ID AAY6726 standard; protein; 425 AA.

XX

AC AAY6726;

XX

DT 26-SEP-2000 (first entry)

XX

DE Human MADr3 polypeptide.

XX

KW MADr3; decapentaplegic; DPP; MAD; Mothers against DPP; vulneryary;

KW transforming growth factor beta; TGF-beta; anti-arthritis; osteopathic;

KW nephrotropic; cardiant; ophthalmological; nootropic; neuroprotective;

KW anti-ulcer; anti-arteriosclerotic; vasotrophic; anti-diabetic.

XX

OS Homo sapiens.

XX

FN WO2000035470-A1.

XX

PD 22-JUN-2000.

XX

PF 15-DEC-1999; 99WO-US029924.

XX

PR 15-DEC-1998; 98US-00212156.

XX

PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX

PI Callahan JF, Laping NJ, Weinstock J;

XX

DR WPI; 2000-442272/38.

XX

DR N-PSDB; AAA51239.

XX

PT Treating diseases and conditions associated with the human protein MADr3

PT e.g. atherosclerosis, renal disease and Alzheimer's disease comprises

PT administration of an inhibitor of ALK5.

XX

PS Disclosure; Page 63-65; 70pp; English.

XX

CC The C-terminal 20 amino acids of this novel MADr3 polypeptide is an

CC inhibitor (antagonist) of ALK5, a receptor involved in the transforming

CC growth factor-beta (TGF-beta) signalling pathway. ALK5 inhibitors can be

CC used to treating diseases and conditions in mammals associated with TGF-

CC beta. MAD (Mothers against DPP) is necessary for signal transduction of

CC DPP (Drosophila decapentaplegic protein). MADr3 is a human homologue of

CC Drosophila MAD. ALK5 inhibitors are useful to treat chronic renal

CC disease, acute renal disease, wound healing, arthritis, osteoporosis,

CC kidney disease, congestive heart failure, impaired neurological function,

CC ocular disorders, trophic conditions, ulcers, Alzheimer's disease,

CC atherosclerosis, restenosis, diabetic neuropathy and any disease where

CC fibrosis is a major component

XX

SQ Sequence 425 AA;

Query Match 100.0%; Score 39; DB 3; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
 Db 179 TPPPGY 184

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TPPPGY 6  
 Db 180 TPPPGY 185

Search completed: October 13, 2005, 15:00:38  
 Job time : 15.5165 secs

RESULT 15  
 AAB80935  
 ID AAB80935 standard; protein; 425 AA.  
 XX AC AAB80935;  
 XX DT 04-JUN-2001 (first entry)  
 XX DE Xenopus Smad3 protein.  
 XX KW Smad; Cytostatic; vulnary; cerebroprotective; immunosuppressive;  
 KW Anti-Hiv; antidiabetic; ophthalmological; antinflammatory; SIM;  
 KW Smad Interaction Motif; tissue repair; fibrotic condition;  
 KW immunosuppression; diabetic nephropathy; tumour.  
 XX OS Xenopus sp.  
 XX PN WO200114413-A2.  
 XX PD 01-MAR-2001.  
 XX PF 25-AUG-2000; 2000WO-GB003265.  
 XX PR 25-AUG-1999; 99GB-00020000.  
 XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX PA Germain SE, Hill CS, Howell MT;  
 XX PI WPI; 2001-265836/27.  
 XX DR N-PSDB; AAF81388.  
 XX Polypeptide capable of interacting with a Smad polypeptide, useful in the  
 PT treatment of cancer and for tissue remodelling or healing of a wound,  
 PT injury or surgery, comprises a Smad Interaction Motif and is less than 32  
 PT amino acids in length.  
 XX Claim 6; Fig 12; 179pp; English.  
 XX The present invention relates to peptides capable of interacting with a  
 CC Smad protein, comprising a Smad Interaction Motif (SIM; amino acid  
 CC sequence PP(T/N)K). The present sequence is a Smad protein. Smad proteins  
 CC are a family of highly conserved, intracellular proteins that signal  
 CC cellular responses downstream of Transforming Growth Factor-beta (TGF-  
 CC beta) family serine/threonine kinase receptors. The SIM is thought to be  
 CC necessary for interaction with the MH2 domain of Smad2. The Smad-  
 CC interacting peptides of the present invention are useful in the  
 CC manufacture of a medicament for the treatment of a patient in need of  
 CC modulation of activin or TGF-beta signalling; cancer; a patient in need  
 CC of reducing extracellular matrix deposition, encouraging tissue repair  
 CC and/or regeneration, tissue remodelling or healing of a wound, injury or  
 CC surgery, or reducing scar tissue formation arising from injury to the  
 CC brain, a patient with or at risk of end-stage organ failure, pathologic  
 CC extracellular matrix accumulation, a fibrotic condition, disease states  
 CC associated with immunosuppression (such as different forms of malignancy,  
 CC chronic degenerative diseases, and AIDS), diabetic nephropathy, tumour  
 CC growth, kidney damage (for e.g. obstructive neuropathy, IGA nephropathy  
 CC or non-inflammatory renal disease) or renal fibrosis  
 XX Sequence 425 AA;

Query Match 100.0%; Score 39; DB 4; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:50:53 ; Search time 2.85714 Seconds  
(without alignments)  
156.763 Million cell updates/sec

Title: US-09-385-918-18

Perfect score: 39

Sequence: 1 TPPPGY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	16	2	US-08-844-312-13
2	39	100.0	16	2	US-08-844-312-14
3	39	100.0	424	3	US-09-087-134-8
4	39	100.0	425	2	US-08-732-028-2
5	39	100.0	425	3	US-09-096-776B-8
6	39	100.0	425	3	US-09-183-228-2
7	39	100.0	425	4	US-09-923-922-8
8	39	100.0	425	4	US-09-949-016-6670
9	39	100.0	436	3	US-08-840-767-6
10	39	100.0	451	4	US-09-949-016-7630
11	39	100.0	467	3	US-08-701-582D-2
12	39	100.0	467	3	US-08-701-582D-4
13	39	100.0	467	3	US-09-082-039A-2
14	39	100.0	467	3	US-09-082-039A-15
15	39	100.0	467	3	US-08-840-767-42
16	39	100.0	467	3	US-08-840-767-50
17	39	100.0	467	3	US-09-096-776B-7
18	39	100.0	467	3	US-09-087-134-2
19	39	100.0	467	3	US-09-087-134-5
20	39	100.0	467	3	US-09-552-138-2
21	39	100.0	467	3	US-09-552-138-15
22	39	100.0	467	4	US-08-580-031A-15
23	39	100.0	467	4	US-09-923-922-7
24	39	100.0	477	3	US-09-113-309-3
25	39	100.0	477	3	US-09-521-109-3
26	39	100.0	477	3	US-08-840-767-2
27	39	100.0	477	3	US-09-562-332-3

28 39 100.0 485 4 US-09-949-016-7633 Sequence 7633, Ap  
29 39 100.0 718 4 US-09-417-197-75 Sequence 75, Appl  
30 39 100.0 719 4 US-09-417-197-51 Sequence 51, Appl  
31 35 89.7 619 1 US-08-472-934-4 Sequence 4, Appl  
32 35 89.7 619 1 US-08-472-934-12 Sequence 12, Appl  
33 35 89.7 619 2 US-08-323-460A-4 Sequence 4, Appl  
34 35 89.7 619 2 US-08-461-146C-4 Sequence 4, Appl  
35 35 89.7 619 2 US-08-461-146C-12 Sequence 12, Appl  
36 35 89.7 619 3 US-08-461-145C-4 Sequence 4, Appl  
37 35 89.7 619 3 US-08-461-145C-12 Sequence 12, Appl  
38 35 89.7 619 3 US-09-423-890-4 Sequence 4, Appl  
39 35 89.7 619 3 US-09-423-890-10 Sequence 10, Appl  
40 35 89.7 619 3 US-08-628-829-6 Sequence 6, Appl  
41 35 89.7 619 3 US-08-628-829-8 Sequence 8, Appl  
42 35 89.7 624 4 US-09-949-016-10942 Sequence 10942, A  
43 35 89.7 626 1 US-08-472-934-6 Sequence 6, Appl  
44 35 89.7 626 2 US-08-323-460A-6 Sequence 6, Appl  
45 35 89.7 626 2 US-08-461-146C-6 Sequence 6, Appl

## ALIGNMENTS

### RESULT 1

US-08-844-312-13  
; Sequence 13, Application US/08844312  
; Patent No. 5948639  
; GENERAL INFORMATION:  
; APPLICANT: Carlos J. Gimeno and Dean A. Falb  
; TITLE OF INVENTION: No. 5948639el TGF-(SYMBOL 98 \f "Symbol") Pathway Genes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/844.312  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silveri, Jean M.  
; REGISTRATION NUMBER: 39,030  
; REFERENCE/DOCKET NUMBER: MNI-013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-844-312-13

Query Match 100.0%; Score 39; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6

Db 7 TPPPGY 12

```
RESULT 2
US-08-844-312-14
; Sequence 14, Application US/08844312
; Patent No. 5948539
; GENERAL INFORMATION:
; APPLICANT: Carlos J. Gimeno and Dean A. Falb
; TITLE OF INVENTION: No. 5948639e1 TGF-{SYMBOL 98 \f "Symbol"} Pathway Genes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/844,312
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-844-312-14

Query Match 100.0%; Score 39; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPGY 6
Db 7 TPPPGY 12

RESULT 3
US-09-087-134-8
; Sequence 8, Application US/09087134
; Patent No. 6365711
; GENERAL INFORMATION:
; APPLICANT: Malcolm Whitman and Xin Chen
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: TGF-BETA SUPERFAMILY SIGNALLING
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,134
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Schreck, Patricia A.
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: SBC ATG 50030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,134
; FILING DATE: 27-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,991
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bicker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 00246/501002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...1
; OTHER INFORMATION: Human Smad3
; US-09-087-134-8

Query Match 100.0%; Score 39; DB 3; Length 424;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPPPGY 6
Db 178 TPPPGY 183

RESULT 4
US-08-732-028-2
; Sequence 2, Application US/08732028
; Patent No. 5868693
; GENERAL INFORMATION:
; APPLICANT: Laping, Nicholas J.
; TITLE OF INVENTION: Human MAD Protein and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,028
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Schreck, Patricia A.
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: SBC ATG 50030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-732-028-2

Query Match      100.0%; Score 39; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      179 TTPPGY 184

RESULT 5
US-09-096-776B-8
; Sequence 8, Application US/09096776B
; Patent No. 6270994
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohei
; APPLICANT: Kawabata, Masahiro
; TITLE OF INVENTION: SMAD6 AND USES THEREOF
; FILE REFERENCE: L0461/7038
; CURRENT APPLICATION NUMBER: US/09/096,776B
; CURRENT FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: US 60/049,990
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: US 60/053,040
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: US 60/066,173
; PRIOR FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-096-776B-8

Query Match      100.0%; Score 39; DB 3; Length 425;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      179 TTPPGY 184

RESULT 6
US-09-183-228-2
; Sequence 2, Application US/09183228
; Patent No. 6309856
; GENERAL INFORMATION:
; APPLICANT: LAPING, NICHOLAS J.
; TITLE OF INVENTION: HUMAN MAD PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,228
; FILING DATE: 30-OCT-1998
```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/732,028
; FILING DATE: 16-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: ATG-50030-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-183-228-2

Query Match      100.0%; Score 39; DB 3; Length 425;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      179 TTPPGY 184

RESULT 7
US-09-923-922-8
; Sequence 8, Application US/09923922
; Patent No. 6534476
; GENERAL INFORMATION:
; APPLICANT: Miyazono, Kohei
; APPLICANT: Kawabata, Masahiro
; TITLE OF INVENTION: SMAD6 AND USES THEREOF
; FILE REFERENCE: L0461/7120
; CURRENT APPLICATION NUMBER: US/09/923,922
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 09/096,776
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: US 60/049,990
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: US 60/053,040
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: US 60/066,173
; PRIOR FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-922-8

Query Match      100.0%; Score 39; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      179 TTPPGY 184

RESULT 8
US-09-949-016-6670
; Sequence 6670, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6670
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6670

Query Match          100.0%; Score 39; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 179 TTPPGY 184

RESULT 9
US-08-840-767-6
; Sequence 6, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggins, Gregory J.
; APPLICANT: Thiagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-840-767-6

Query Match          100.0%; Score 39; DB 3; Length 436;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 190 TTPPGY 195

RESULT 10
US-09-949-016-7630
; Sequence 7630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7630
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7630

Query Match          100.0%; Score 39; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 205 TTPPGY 210

RESULT 11
US-08-701-582D-2
; Sequence 2, Application US/08701582D
; Patent No. 6017755
; GENERAL INFORMATION:
; APPLICANT: WRANA, Jeffrey
; APPLICANT: ATTISANO, Lilliana
; APPLICANT: SCHERER, Stephen W.
; TITLE OF INVENTION: MADR2 TUMOUR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701.582D
; FILING DATE: 22-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 024916-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-582D-2

Query Match          100.0%; Score 39; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 220 TTPPGY 225

RESULT 12
US-08-701-582D-4
```

; Sequence 4, Application US/08701582D  
; Patent No. 6017755  
; GENERAL INFORMATION:  
; APPLICANT: WRANA, Jeffrey  
; APPLICANT: ATTISANO, Liliana  
; APPLICANT: SCHERER, Stephen W.  
; TITLE OF INVENTION: MADR2 TUMOUR SUPPRESSOR GENE  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Deane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/701.582D  
; FILING DATE: 22-AUG-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REA, Teresa Stanek  
; REGISTRATION NUMBER: 30,427  
; REFERENCE/DOCKET NUMBER: 024916-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 467 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-701-582D-4

Query Match 100.0%; Score 39; DB 3; Length 467;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 220 TTPPGY 225

RESULT 13  
US-09-082-039A-2  
; Sequence 2, Application US/09082039A  
; Patent No. 6103869  
; GENERAL INFORMATION:  
; APPLICANT: Souhelnytskyi, Serheyi  
; APPLICANT: Tamaki, Kiyoshi  
; APPLICANT: Engstr m, Ulla  
; APPLICANT: Wernstedt, Christer  
; APPLICANT: Piek, Esther  
; APPLICANT: ten Dijke, Peter  
; APPLICANT: Heldin, Carl-Henrik  
; TITLE OF INVENTION: SMAD2 PHOSPHORYLATION AND INTERACTION  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.039A  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/047,807

; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.039A  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/047,807  
; FILING DATE: 20-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/081,313  
; FILING DATE: 10-APR-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: John R. Van Amsterdam  
; REGISTRATION NUMBER: 40,212  
; REFERENCE/DOCKET NUMBER: L0461/7036  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 467 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-09-082-039A-2

Query Match 100.0%; Score 39; DB 3; Length 467;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 220 TTPPGY 225

RESULT 14  
US-09-082-039A-15  
; Sequence 15, Application US/09082039A  
; Patent No. 6103869  
; GENERAL INFORMATION:  
; APPLICANT: Souhelnytskyi, Serheyi  
; APPLICANT: Tamaki, Kiyoshi  
; APPLICANT: Engstr m, Ulla  
; APPLICANT: Wernstedt, Christer  
; APPLICANT: Piek, Esther  
; APPLICANT: ten Dijke, Peter  
; APPLICANT: Heldin, Carl-Henrik  
; TITLE OF INVENTION: SMAD2 PHOSPHORYLATION AND INTERACTION  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.039A  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/047,807

```
; FILING DATE: 20-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/081,313
; FILING DATE: 10-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: John R. Van Amsterdam
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FRAGMENT TYPE: internal
;
US-09-082-039A-15

Query Match 100.0%; Score 39; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TPPPGY 6
Db 220 TPPPGY 225
```

```
RESULT 15
US-08-840-767-42
; Sequence 42, Application US/08840767B
; Patent No. 6255464
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Riggs, Gregory J.
; APPLICANT: Thagalingam, Sam
; TITLE OF INVENTION: MAD-Related Genes in the Human
; FILE REFERENCE: 01107.05548
; CURRENT APPLICATION NUMBER: US/08/840,767B
; CURRENT FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: 60/015,823
; EARLIER FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-840-767-42
```

```
Query Match 100.0%; Score 39; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TPPPGY 6
Db 220 TPPPGY 225
```

Search completed: October 13, 2005, 14:03:08  
Job time : 3.98214 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 13:59:34 ; Search time 12.7912 Seconds  
(without alignments)  
195.471 Million cell updates/sec

Title: US-09-385-918-18

Perfect score: 39

Sequence: 1 TPPPGY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	6	14 US-10-307-956-18	Sequence 18, Appl
2	39	100.0	11	14 US-10-307-956-19	Sequence 19, Appl
3	39	100.0	14	14 US-10-307-956-24	Sequence 24, Appl
4	39	100.0	14	14 US-10-307-956-25	Sequence 25, Appl
5	39	100.0	65	16 US-10-425-115-344905	Sequence 344905,
6	39	100.0	72	16 US-10-425-115-353464	Sequence 353464,
7	39	100.0	78	16 US-10-767-701-43198	Sequence 43198, A
8	39	100.0	116	16 US-10-425-115-353466	Sequence 353466,
9	39	100.0	118	16 US-10-767-701-34928	Sequence 34928, A
10	39	100.0	123	16 US-10-425-115-296400	Sequence 296400,
11	39	100.0	132	16 US-10-425-115-353469	Sequence 353469,

12	39	100.0	158	16	US-10-425-115-314821	Sequence 314821,
13	39	100.0	208	16	US-10-425-115-333728	Sequence 333728,
14	39	100.0	227	16	US-10-425-115-242437	Sequence 242437,
15	39	100.0	302	16	US-10-425-115-353473	Sequence 353473,
16	39	100.0	310	15	US-10-425-114-62416	Sequence 62416, A
17	39	100.0	315	15	US-10-425-114-70132	Sequence 70132, A
18	39	100.0	323	16	US-10-425-115-295775	Sequence 295775,
19	39	100.0	419	16	US-10-425-115-331883	Sequence 331883,
20	39	100.0	424	13	US-10-044-442-8	Sequence 8, Appli
21	39	100.0	425	9	US-09-930-317-2	Sequence 2, Appli
22	39	100.0	425	9	US-09-923-922-8	Sequence 8, Appli
23	39	100.0	425	10	US-09-601-534-5	Sequence 5, Appli
24	39	100.0	425	14	US-10-299-886-2	Sequence 2, Appli
25	39	100.0	425	14	US-10-308-279-46	Sequence 46, Appli
26	39	100.0	425	15	US-10-390-553-8	Sequence 8, Appli
27	39	100.0	425	16	US-10-648-593-198	Sequence 198, App
28	39	100.0	444	15	US-10-369-493-23166	Sequence 23166, A
29	39	100.0	467	9	US-09-923-922-7	Sequence 7, Appli
30	39	100.0	467	10	US-09-601-534-4	Sequence 4, Appli
31	39	100.0	467	13	US-10-095-492-15	Sequence 15, Appli
32	39	100.0	467	13	US-10-119-099-2	Sequence 2, Appli
33	39	100.0	467	13	US-10-119-099-15	Sequence 15, Appli
34	39	100.0	467	13	US-10-044-442-2	Sequence 2, Appli
35	39	100.0	467	13	US-10-044-442-5	Sequence 5, Appli
36	39	100.0	467	15	US-10-390-553-7	Sequence 7, Appli
37	39	100.0	486	20	US-11-097-143-3774	Sequence 3774, Ap
38	39	100.0	664	16	US-10-425-115-253413	Sequence 253413,
39	39	100.0	718	14	US-10-072-036-75	Sequence 75, Appl
40	39	100.0	719	14	US-10-072-036-51	Sequence 51, Appl
41	39	100.0	763	16	US-10-425-115-331129	Sequence 331129,
42	39	100.0	848	16	US-10-425-115-295770	Sequence 295770,
43	39	100.0	961	16	US-10-425-115-331333	Sequence 331333,
44	39	100.0	1394	16	US-10-437-963-108833	Sequence 108833,
45	39	100.0	1470	17	US-10-732-923-1639	Sequence 1639, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-307-956-18  
; Sequence 18, Application US/10307956  
; Publication No. US20030119072A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoeckstra, Merl F.  
; APPLICANT: Xie, Weilin  
; APPLICANT: Murray, Brion  
; APPLICANT: Mercurio, Frank  
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL  
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS  
; FILE REFERENCE: 860098.433  
; CURRENT APPLICATION NUMBER: US/10/307,956  
; CURRENT FILING DATE: 2002-12-02  
; PRIOR APPLICATION NUMBER: US/09/385,918  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-307-956-18

Query Match 100.0%; Score 39; DB 14; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6

Db 1 TPPPGY 6

#### RESULT 2

```
US-10-307-956-19
; Sequence 19, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = Isoleucine or Leucine
US-10-307-956-19

Query Match      100.0%; Score 39; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      1 TTPPGY 6

RESULT 3
US-10-307-956-24
; Sequence 24, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-24

Query Match      100.0%; Score 39; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      4 TTPPGY 9

RESULT 4
US-10-307-956-25
; Sequence 25, Application US/10307956
; Publication No. US20030119072A1
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Xie, Weilin
; APPLICANT: Murray, Brion
; APPLICANT: Mercurio, Frank
; TITLE OF INVENTION: METHODS FOR MODULATING SIGNAL
; TITLE OF INVENTION: TRANSDUCTION MEDIATED BY TGF-BETA AND RELATED PROTEINS
; FILE REFERENCE: 860098.433
; CURRENT APPLICATION NUMBER: US/10/307,956
; CURRENT FILING DATE: 2002-12-02
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US/09/385,918
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-307-956-25

Query Match      100.0%; Score 39; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      4 TTPPGY 9

RESULT 5
US-10-425-115-344905
; Sequence 344905, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 344905
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(65)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77720C.1.pep
US-10-425-115-344905

Query Match      100.0%; Score 39; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
Db      1 TTPPGY 6

RESULT 6
US-10-425-115-353464
; Sequence 353464, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```



; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 353464  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_85534C.1.pep  
US-10-425-115-353464

Query Match 100.0%; Score 39; DB 16; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 63 TTPPGY 68  
|||||

## RESULT 7

US-10-767-701-43198  
; Sequence 43198, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 43198  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7378\_1.pep  
US-10-767-701-43198

Query Match 100.0%; Score 39; DB 16; Length 78;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 69 TTPPGY 74  
|||||

## RESULT 8

US-10-425-115-353466  
; Sequence 353466, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 353466  
; LENGTH: 116  
; TYPE: PRT

; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_85536C.1.pep  
US-10-425-115-353466

Query Match 100.0%; Score 39; DB 16; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 107 TTPPGY 112  
|||||

## RESULT 9

US-10-767-701-34928  
; Sequence 34928, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 34928  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CS6676\_1.pep  
US-10-767-701-34928

Query Match 100.0%; Score 39; DB 16; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
Db 7 TTPPGY 12  
|||||

## RESULT 10

US-10-425-115-296400  
; Sequence 296400, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 296400  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_33391C.1.pep  
US-10-425-115-296400

Query Match 100.0%; Score 39; DB 16; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6

```
Db      34 TTPPGY 39
|||||
RESULT 11
US-10-425-115-353469
; Sequence 353469, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 353469
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_85539C.1.pep
US-10-425-115-353469

Query Match      100.0%; Score 39; DB 16; Length 132;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
|||||
Db      123 TTPPGY 128

RESULT 12
US-10-425-115-314821
; Sequence 314821, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 314821
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_50179C.1.pep
US-10-425-115-314821

Query Match      100.0%; Score 39; DB 16; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
|||||
Db      54 TTPPGY 59

RESULT 13
US-10-425-115-333728
; Sequence 333728, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 333728
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(208)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_67472C.1.pep
US-10-425-115-333728

Query Match      100.0%; Score 39; DB 16; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
|||||
Db      54 TTPPGY 59

RESULT 14
US-10-425-115-242437
; Sequence 242437, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 242437
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(227)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_152689C.1.pep
US-10-425-115-242437

Query Match      100.0%; Score 39; DB 16; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTPPGY 6
|||||
Db      54 TTPPGY 59

RESULT 15
US-10-425-115-353473
; Sequence 353473, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 353473
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_85542C.1.pbp
US-10-425-115-353473

Query Match      100.0%; Score 39; DB 16; Length 302;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TPPEGY 6
        |||||
Db      293 TPPEGY 298

```

Search completed: October 13, 2005, 14:50:03  
Job time : 12.7912 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2005, 14:36:01 ; Search time 2.50549 Seconds  
(without alignments)  
230.414 Million cell updates/sec

Title: US-09-385-918-18

Perfect score: 39

Sequence: 1 TPPPGY 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	274	2 D71443	hypothetical prote
2	39	100.0	424	2 S71798	MAD-3 protein homo
3	39	100.0	444	1 P69904	adenosylmethionine
4	39	100.0	467	2 S71797	MAD-2 protein homo
5	39	100.0	484	2 JEO341	R-SMAD protein, Sm
6	39	100.0	599	2 T10798	phosphorin-S - Vo
7	39	100.0	1101	2 S15271	endoglucanase C (E
8	39	100.0	1711	2 T21432	hypothetical prote
9	35	89.7	151	2 S33621	ADR11-2 protein -
10	35	88.7	347	2 T14313	hypothetical prote
11	35	89.7	376	2 S71558	probable cell wall
12	35	89.7	532	2 T30764	hypothetical prote
13	35	89.7	574	2 AD1863	cobalamin biosynth
14	35	89.7	621	2 T15046	arginine decarboxy
15	35	89.7	711	2 T12525	hypothetical prote
16	35	89.7	2242	2 A57541	pyrimidine synthas
17	34	87.2	107	2 S52507	probable membrane
18	34	87.2	107	2 S18529	hypothetical prote
19	34	87.2	136	2 T46900	hypothetical prote
20	34	87.2	139	2 C87544	hypothetical prote
21	34	87.2	141	2 S09804	hypothetical prote
22	34	87.2	156	2 F72506	hypothetical prote
23	34	87.2	198	2 E86261	Fl3K23.6 protein -
24	34	87.2	199	2 S14981	extensin class I (
25	34	87.2	232	2 D96663	unknown protein, 5
26	34	87.2	241	2 G70578	hypothetical prote
27	34	87.2	260	2 AB2080	hypothetical prote
28	34	87.2	265	2 T36192	hypothetical prote
29	34	87.2	276	2 G95887	probable ABC trans

30	34	87.2	291	2 G84494	hypothetical prote
31	34	87.2	308	1 A34082	branched-chain-ami
32	34	87.2	309	2 AD0924	probable membrane
33	34	87.2	319	2 T36845	hypothetical prote
34	34	87.2	319	2 T34525	hypothetical prote
35	34	87.2	325	2 D70728	hypothetical prote
36	34	87.2	344	2 T33421	hypothetical prote
37	34	87.2	358	2 JC4311	CCAAT/enhancer bin
38	34	87.2	358	2 A54265	poly(beta-d-mannur
39	34	87.2	367	2 H83202	hypothetical prote
40	34	87.2	367	2 T32730	poly(beta-D-mannur
41	34	87.2	368	2 JN0777	hypothetical prote
42	34	87.2	380	2 T48052	hypothetical prote
43	34	87.2	393	2 T49578	hypothetical prote
44	34	87.2	395	2 I49575	CCAAT/enhancer bin
45	34	87.2	399	2 AD1808	hypothetical prote

ALIGNMENTS

RESULT 1

D71443

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004

C:Accession: D71443

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech

erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans

C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: D71443

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-274 <BE>

A:Cross-references: UNIPROT:O23583; GB:Z97343; NID:g2245073; PID:e327059; PID:g2245099

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 100.0%; Score 39; DB 2; Length 274;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6

Db 196 TPPPGY 201

RESULT 2

S71798

MAD-3 protein homolog - human

C:Species: Homo sapiens (man)

C:Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 02-Sep-2000

C:Accession: S71798

R:Zhang, Y.; Feng, X.H.; Wu, R.Y.; Derynck, R.

Nature 383, 168-172, 1996

A:Title: Receptor-associated Mad homologues synergize as effectors of the TGF-beta response

A:Reference number: S71797; MUID:96371046; PMID:8774881

A:Accession: S71798

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-424 <ZHA>

A:Cross-references: EMBL:U68019

C:Superfamily: human transcription regulator MAD-4

Query Match 100.0%; Score 39; DB 2; Length 424;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 Db 178 TTPPGY 183

RESULT 3  
 F69904  
 adenosylmethionine-8-amino-7-oxononanoate homolog yodT - Bacillus subtilis  
 N:Alternate names: probable transaminase (EC 2.6.1.1-)  
 C:Species: Bacillus subtilis  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: F69904  
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel  
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F69904  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-444 <KUN>  
 A:Cross-references: UNIPROT:O34662; GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13865.  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yodT  
 C:Superfamily: beta-alanine-pyruvate transaminase  
 C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate  
 F:268/Banding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 100.0%; Score 39; DB 1; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 Db 216 TTPPGY 221

RESULT 4  
 S71797  
 MAD-2 protein homolog - human  
 C:Species: Homo sapiens (man)  
 C>Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
 C:Accession: S71797  
 R:Zhang, Y.; Feng, X.H.; Wu, R.Y.; Derynck, R.  
 Nature 383, 168-172, 1996  
 A:Title: Receptor-associated Mad homologues synergize as effectors of the TGF-beta resp  
 A:Reference number: S71797; MUID:96371046; PMID:8774881  
 A:Accession: S71797  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-467 <ZHA>  
 A:Cross-references: UNIPROT:Q15796; EMBL:U68018; NID:g1552529; PIDN:AAB17087.1; PID:g155  
 C:Superfamily: human transcription regulator MAD-4

Query Match 100.0%; Score 39; DB 2; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6

Db 220 TTPPGY 225

RESULT 5  
 JE0341  
 R-SMAD protein, Smox - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:Accession: JE0341  
 R:D.Henderson, K.; J.Andrew, D.  
 Biochem. Biophys. Res. Commun. 252, 195-201, 1998  
 A:Title: Identification of a novel Drosophila SMAD on the X chromosome.  
 A:Reference number: JE0341; MUID:99032828; PMID:9813169  
 A:Accession: JE0341  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-484 <DAH>  
 A:Cross-references: UNIPROT:Q9TZQ2; GB:AF078529; NID:g3978423; PIDN:AAC83344.1; PID:g397  
 C:Genetics:  
 A:Cross-references: FlyBase:FBgn0025800  
 A:Superfamily: human transcription regulator MAD-4

Query Match 100.0%; Score 39; DB 2; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 Db 250 TTPPGY 255

RESULT 6  
 TI0798  
 pherophorin-S - Volvox carteri  
 C:Species: Volvox carteri  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:Accession: TI0798  
 R:Godl, K.; Hallmann, A.; Wenzl, S.; Sumper, M.  
 EMBO J. 16, 25-34, 1997  
 A:Title: Differential targeting of closely related ECM-glycoproteins: The pherophorin f  
 A:Reference number: Z17154; MUID:97162277; PMID:9009264  
 A:Accession: TI0798  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-599 <GD>  
 A:Cross-references: UNIPROT:P93797; EMBL:Y07752; NID:g1655698; PIDN:CAA69032.1; PID:g165  
 A:Experimental source: strain HK 10; sub.species Nagariensis  
 A>Note: in contrast to the other pherophorins, pherophorin-S is targeted to the cell-fre  
 C:Keywords: extracellular matrix; glycoprotein; pheromone

Query Match 100.0%; Score 39; DB 2; Length 599;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6  
 Db 180 TTPPGY 185

RESULT 7  
 SI5271  
 endoglucanase C (EC 3.2.1.-) - Cellulomonas fimi  
 C:Species: Cellulomonas fimi  
 C>Date: 21-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
 C:Accession: SI5271; A43636  
 R:Coutinho, J.B.; Moser, B.; Kilburn, D.G.; Warren, R.A.J.; Miller Jr., R.C.  
 Mol. Microbiol. 5, 1221-1233, 1991  
 A:Title: Nucleotide sequence of the endoglucanase C gene (cenc) of Cellulomonas fimi, it  
 A:Reference number: SI5271; MUID:92065819; PMID:1956299  
 A:Accession: SI5271  
 A:Status: preliminary  
 A:Molecule type: DNA

A;Residues: 1-1101 <COU>  
A;Cross-references: UNIPROT:P14090; EMBL:X57958; NID:g48847; PIDN:CAA40993.1; PID:g58098  
A;Note: the authors translated the codon GAC for residues 361, 380, 400, 550, 670, and 8  
R;Moser, B.; Gilkes, N.R.; Kilburn, D.G.; Warren, R.A.J.; Miller Jr., R.C.  
Appl. Environ. Microbiol. 55, 2480-2487, 1989  
A;Title: Purification and characterization of endoglucanase C of Cellulomonas fimi, clone  
A;Reference number: A43636; MUID:90103465; PMID:2604391  
A;Accession: A43636  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-64, 'w' <MOS>  
A;Cross-references: GB:M29707; NID:g144417; PIDN:AAA23087.1; PID:g144420  
C;Genetics:  
A;Gene: cenC  
A;Start codon: GTG  
C;Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 39; DB 2; Length 1101;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
|||||  
Db 338 TPPPGY 333

RESULT 8  
T21432  
hypothetical protein F26H11.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T21432  
R;Barlow, K.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19421  
A;Accession: T21432  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1711 <WIL>  
A;Cross-references: UNIPROT:Q45409; EMBL:Z81515; PIDN:CAB04197.1; GSPDB:GN000020; CESP:F2  
A;Experimental source: clone F26H11  
C;Genetics:  
A;Gene: CESP:F26H11.2  
A;Map position: 2  
A;Introns: 56/3; 154/2; 373/2; 417/3; 505/2; 596/1; 703/2; 847/3; 922/3; 1034/2; 1078/2;

Query Match 100.0%; Score 39; DB 2; Length 1711;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
|||||  
Db 1336 TPPPGY 1341

RESULT 9  
S33621  
ADRI1-2 protein - soybean (fragment)  
C;Species: Glycine max (soybean)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S33621  
R;Datta, N.; LaFayette, P.R.; Kroner, P.A.; Nagao, R.T.; Key, J.L.  
Plant Mol. Biol. 21, 859-869, 1993  
A;Title: Isolation and characterization of three families of auxin down-regulated cDNA c  
A;Reference number: S33620; MUID:93222483; PMID:8096772  
A;Accession: S33621  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-151 <DAT>  
A;Cross-references: UNIPROT:Q06763; EMBL:X69640; NID:g296442; PIDN:CAA49341.1; PID:g2964  
C;Genetics:  
A;Gene: ADRI1  
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 89.7%; Score 35; DB 2; Length 151;  
Best Local Similarity 83.3%; Pred. No. 59;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
|||||  
Db 142 SPPEGY 147

## RESULT 10

T14313

hypothetical protein - carrot

C;Species: Daucus carota (carrot)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T14313

R;Kawahara, R.; Matsumoto, M.; Ozeki, Y.; Toya, T.; Ito, M.; Fujiwata, A.; Komamine, A.

submitted to the EMBL Data Library, January 1997

A;Description: Isolation and characterization of organ specific genes in carrot.

A;Reference number: Z17973

A;Accession: T14313

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-347 &lt;KAW&gt;

A;Cross-references: UNIPROT:P93705; EMBL:AB000505

A;Experimental source: cultivar Kurodagosun; hypocotyl

C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 89.7%; Score 35; DB 2; Length 347;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
|||||  
Db 338 TPPPGF 343

## RESULT 11

S71558

probable cell wall-plasma membrane linker protein PRP precursor - rape

N;Alternate names: hybrid-proline-rich protein

C;Species: Brassica napus (rape)

C;Date: 27-Nov-1997 #sequence\_revision 12-Dec-1997 #text\_change 09-Jul-2004

C;Accession: S71558

R;Goodwin, W.; Pallas, J.A.; Jenkins, G.I.

Plant Mol. Biol. 31, 771-781, 1996

A;Title: Transcripts of a gene encoding a putative cell wall-plasma membrane linker prot

A;Reference number: S71558; MUID:96400032; PMID:8806408

A;Accession: S71558

A;Molecule type: DNA

A;Residues: 1-376 &lt;GOO&gt;

A;Cross-references: UNIPROT:Q39353; EMBL:X94976; NID:g1155067; PIDN:CAA64425.1; PID:g115

C;Superfamily: hydroxyproline-rich glycoprotein

F;1-27/Domain: signal sequence #status predicted &lt;SIG&gt;

F;28-376/Product: probable cell wall-plasma membrane linker protein PRP #status predicted

Query Match 89.7%; Score 35; DB 2; Length 376;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6  
|||||  
Db 367 TPPPGF 372

## RESULT 12

T30764

hypothetical protein 162R - Mollusca contagiosum virus 1

N;Alternate names: MC162R

C;Species: Mollusca contagiosum virus 1

C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C;Accession: T30764

R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.

Science 273, 813-816, 1996

A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re

A;Reference number: Z20876; MUID:96325459; PMID:8670425

A;Accession: T30764

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-532 <SEN>

A;Cross-references: UNIPROT:Q98328; EMBL:U60315; PIDN:AAC55290.1

C;Genetics:

A;Note: MCI62R

Query Match

Best Local Similarity 89.7%; Score 35; DB 2; Length 532;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPPPGY 6

121 TPPPGF 126

RESULT 13

AD1863

cobalamin biosynthesis precorrin-3 methylase [imported] - Nostoc sp. (strain PCC 7120)

A;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 16-Aug-2004

C;Accession: AD1863

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi

Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AD1863

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-574 <KUR>

A;Cross-references: UNIPROT:Q8YZK4; GB:BA000019; PIDN:BAB72411.1; PID:g17129798; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all0453

C;Superfamily: Bifunctional CblG/precorrin methyltransferase

Query Match

Best Local Similarity 89.7%; Score 35; DB 2; Length 574;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPPPGY 6

139 SPPPGY 144

RESULT 14

T15046

arginine decarboxylase (EC 4.1.1.19) ADC-1 - wood tobacco

C;Species: Nicotiana sylvestris (wood tobacco)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T15046

R;Nakakita, M.

submitted to the EMBL Data Library, April 1998

A;Description: Putative cDNA for arginine decarboxylase from Nicotiana sylvestris.

A;Reference number: Z18277

A;Accession: T15046

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-621 <NAK>

A;Cross-references: UNIPROT:O64453; EMBL:AB012873

C;Genetics:

A;Gene: ADC-1

C;Function:

A;Description: catalyzes decarboxylation of arginine into agmatine

A;Pathway: putrescine biosynthesis

C;Superfamily: arginine decarboxylase

C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match

Best Local Similarity 89.7%; Score 35; DB 2; Length 621;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPPPGY 6

13 SPPPGY 18

RESULT 15

T12525

hypothetical protein DKFZp434L243.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 09-Jul-2004

C;Accession: T12525

R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A;Reference number: Z17524

A;Accession: T12525

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-711 <WAM>

A;Cross-references: UNIPROT:Q9Y4Q9; EMBL:AL080140

A;Experimental source: adult testis; clone DKFZp434L243

C;Genetics:

A;Note: DKFZp434L243.1

Query Match

Best Local Similarity 89.7%; Score 35; DB 2; Length 711;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPPPGY 6

234 SPPPGY 239

Search completed: October 13, 2005, 15:11:27

Job time : 4.50549 secs



Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	39	100.0	77	2	Q8WNY3	Q8WNY3 ovis aries
2	39	100.0	125	2	Q887P0	Q887P0 pseudomonas
3	39	100.0	156	2	Q864T4	Q864T4 ovis aries
4	39	100.0	171	2	Q8MKC6	Q8MKC6 ovis aries
5	39	100.0	184	2	Q9G232	Q9G232 asparagus o
6	39	100.0	219	2	Q8WVX9	Q8WVX9 cucumis sat
7	39	100.0	254	2	Q94EV4	Q94EV4 zea mays (m
8	39	100.0	274	2	O23583	O23583 arabidopsis
9	39	100.0	308	2	Q8GNL2	Q8GNL2 drosophila
10	39	100.0	313	2	Q919P1	Q919P1 gallus gall
11	39	100.0	378	2	Q7Q1I2	Q7Q1I2 anopheles g
12	39	100.0	385	2	Q919P2	Q919P2 gallus gall
13	39	100.0	422	2	Q6QAN7	Q6QAN7 carassius a
14	39	100.0	423	2	O6GHX9	O6GHX9 brachydanio
15	39	100.0	423	2	Q8AY16	Q8AY16 brachydanio
16	39	100.0	425	1	SM33_HUMAN	SM34022 homo sapien
17	39	100.0	425	1	SM33_MOUSE	Q8BUN5 mus musculu
18	39	100.0	425	1	SM33_PIG	SM34024 sus scrofa
19	39	100.0	425	1	SM33_RAT	SM34025 rattus norv
20	39	100.0	425	2	Q90VE5	Q90VE5 xenopus lae
21	39	100.0	425	2	Q68EP5	Q68EP5 xenopus tro
22	39	100.0	425	2	Q8AY15	Q8AY15 brachydanio
23	39	100.0	426	1	SM33_CHICK	SM34023 gallus gall
24	39	100.0	436	2	O6G156	O6G156 streptomyce
25	39	100.0	437	2	Q6VP00	Q6VP00 mus musculu
26	39	100.0	444	1	YODT_BACSU	O34662 bacillus su
27	39	100.0	450	2	Q8T8C9	Q8T8C9 halocynthia
28	39	100.0	450	2	Q7PK45	Q7PK45 anopheles g
29	39	100.0	462	2	Q8U7L2	Q8U7L2 arabidopsis
30	39	100.0	467	1	SM32_HUMAN	Q15796 homo sapien
31	39	100.0	467	1	SM32_MOUSE	Q62432 mus musculu

```

RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Pedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016860; AA054776.1; -.
DR TIGR; PSPT01251; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 125 AA; 13651 MW; 450717203FF6205B CRC64;

Query Match 100.0%; Score 39; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 72 TPPPGY 77

RESULT 3
Q864T4 PRELIMINARY; PRT; 156 AA.
AC Q864T4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Smad2 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12606325;
RA Dupont J., McNeilly J., Vaiman A., Canepa S., Combarnous Y.,
RA Taragnat C.;
RT "Activin Signaling Pathways in Ovine Pituitary and L T2 Gonadotrope
RL Biol. Reprod. 68:1877-1887(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Dupont J.S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AY185300; AA027459.1; -.
DR HSSP; Q15796; IDEV.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR003619; Dwarfin A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
FT NON TER 156
FT NON TER 156
SQ SEQUENCE 156 AA; 17618 MW; 7A020FEBBF938DF3 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 83 TPPPGY 88

RESULT 4
Q8MKC6 PRELIMINARY; PRT; 171 AA.
AC Q8MKC6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Smad3 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza C.J.H., Baird D.T.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF508024; AAM34245.1; -.
DR HSSP; Q15796; IDEV.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03166; MH2; 1.
FT NON TER 171
FT NON TER 171
SQ SEQUENCE 171 AA; 18889 MW; 6E7D496D290BAE62 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 50 TPPPGY 55

RESULT 5
Q96232 PRELIMINARY; PRT; 184 AA.
AC Q96232;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proline-rich-like protein.
OS Asparagus officinalis (Garden asparagus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Asparagaceae;
OC Asparagus.
OX NCBI_TaxID=4686;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spear tips;
RX MEDLINE=96416434; PubMed=8819317;
RA King G.A., O'Donoghue, Borst, Davies, Moyle, Farnden;
RT "Identification and characterization of an mRNA encoding a proline-
RT rich protein that rapidly declines in abundance in the tips of
RT harvested asparagus spears.";
RL Plant Cell Physiol. 37:706-710(1996).
DR EMBL; X82413; CAA57810.1; -.
DR HSSP; P24337; IHYP.
DR InterPro; IPR003612; AAL.
DR Pfam; PF00234; TRYP_alpha_amy1; 1.
DR SMART; SM00499; AAL; 1.
SQ SEQUENCE 184 AA; 19292 MW; BD97DD5F1F25E37D CRC64;

Query Match 100.0%; Score 39; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 175 TPPPGY 180

```

```

RESULT 6
Q8VWX9 PRELIMINARY; PRT; 219 AA.
AC Q8VWX9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Extensin-like protein (Fragment).
GN Name=EXTL;
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Male and female floral buds;
RA Kahana A., Silberstein L., Kessler N., Perl-Treves R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104392; AAL35979.1; -.
DR HSSP; P24337; IHYP.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER 1
SQ SEQUENCE 219 AA; 22299 MW; 152B165EDA4A6F20 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPEGY 6
Db 210 TPPEGY 215

RESULT 7
Q94EV4 PRELIMINARY; PRT; 254 AA.
AC Q94EV4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RIR2 orf3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21664393; PubMed=11792865;
RA Fu H., Zheng Z., Dooner H.K.;
RT "Recombination rates between adjacent genic and retrotransposon
regions in maize vary by 2 orders of magnitude.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:1082-1087(2002).
DR EMBL; AF318088; AAK73107.1; -.
DR InterPro; IPR007321; Transposase_28.
DR Pfam; PF04195; Transposase_28; 1.
SQ SEQUENCE 254 AA; 28236 MW; 2C8CF606583C2BB8 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPEGY 6
Db 54 TPPEGY 59

RESULT 8
Q23583 PRELIMINARY; PRT; 274 AA.
ID Q23583
AC Q23583;

Query Match 100.0%; Score 39; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPEGY 6
Db 196 TPPEGY 201

Query Match 100.0%; Score 39; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPEGY 6
Db 196 TPPEGY 201

RESULT 9
Q86NL2 PRELIMINARY; PRT; 308 AA.
ID Q86NL2;
AC Q86NL2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RES3485P.
GN Name=Smox;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Bröckstein P., Hong L., Agbayani A., Carlson J.,
RA George M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT004858; AA045214.1; -.
DR HSSP; Q15796; IDEV.
DR FlyBase; FBgn0025800; Smox.
DR GO; GO:0016358; P:dendrite morphogenesis; NAS.

```

```

DR GO: GO:0016319; P:mushroom body development; IMP.
DR GO: GO:0007476; P:wing morphogenesis; IMP.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00524; DWB; 1.
SQ SEQUENCE 308 AA; 33197 MW; 16FBA0E476E501CB CRC64;

Query Match 100.0%; Score 39; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 74 TPPPGY 79

RESULT 10
Q919P1 PRELIMINARY; PRT; 313 AA.
AC Q919P1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TGF-beta response effector Smad3 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Szendro P.I., Grunenberg U., Eichele G.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230191; AAF36971.1; -.
DR HSSP; Q15796; IDEV.
DR TRANSFAC; T04248; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR003619; Dwarfin.A.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00523; DWA; 1.
DR SMART; SM00524; DWB; 1.
FT NON_TER 1
FT NON_TER 313
SQ SEQUENCE 313 AA; 35476 MW; 0A1FB4BAAA7CF591 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 95 TPPPGY 100

RESULT 11
Q7Q1I2 PRELIMINARY; PRT; 378 AA.
AC Q7Q1I2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP8800.
GN Name=agCG55075; ORFNames=ENSAANGG00000013217;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]

```

---

```

RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008980; EAA13835.1; -.
DR HSSP; Q15796; IDEV.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR003619; Dwarfin.A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
SQ SEQUENCE 378 AA; 42530 MW; AEF8ED24C03777F1 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 142 TPPPGY 147

RESULT 12
Q919P2 PRELIMINARY; PRT; 385 AA.
AC Q919P2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TGF effector Smad2 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Szendro P.I., Eichele G.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230190; AAF36969.1; -.
DR HSSP; Q15796; ICHK.
DR TRANSFAC; T04247; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; Dwarfin.
DR InterPro; IPR003619; Dwarfin.A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00523; DWA; 1.
DR SMART; SM00524; DWB; 1.
FT NON_TER 1
FT NON_TER 385
SQ SEQUENCE 385 AA; 43554 MW; 1A5155BE54DE20C2 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPPPGY 6
Db 143 TPPPGY 148

RESULT 13
Q6QAN7 PRELIMINARY; PRT; 422 AA.
AC Q6QAN7;

```

```

DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Smad3.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
[1]
RN SEQUENCE FROM N.A.
RA Lau M.T., Ge W.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY550113; AAS57862.1; -.
DR HSPF; Q15796; IKH.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; DwarfIn.
DR InterPro; IPR003619; DwarfIn.A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00523; DWA; 1.
DR SMART; SM00524; DWB; 1.
SQ SEQUENCE 422 AA; 47751 MW; 043A37DACADE1F2D CRC64;

Query Match 100.0%; Score 39; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 176 TTPPGY 181

RESULT 14
O66HX9 PRELIMINARY; PRT; 423 AA.
ID Q66HX9
AC Q66HX9;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE MAD, mothers against decapentaplegic homolog 3b.
GN Name=smad3b;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Grumwood J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

RN SEQUENCE FROM N.A.
RA TISSUE=Whole;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC081628; AAH81628.1; -.
DR InterPro; IPR001132; DwarfIn.
DR InterPro; IPR003619; DwarfIn.A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00523; DWA; 1.
DR SMART; SM00524; DWB; 1.
SQ SEQUENCE 423 AA; 47733 MW; DB9092893AC77471 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 177 TTPPGY 182

RESULT 15
Q8AY16 PRELIMINARY; PRT; 423 AA.
ID Q8AY16
AC Q8AY16;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Smad3b.
GN Name=smad3b; Synonyms=madh3b;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RA Pogoda H.-M., Meyer D.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY134490; AAN08604.1; -.
DR HSPF; Q15796; IDEV.
DR ZFIN; ZDB-GENE-030128-4; smad3b.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001132; DwarfIn.
DR InterPro; IPR003619; DwarfIn.A.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF03165; MH1; 1.
DR Pfam; PF03166; MH2; 1.
DR SMART; SM00523; DWA; 1.
DR SMART; SM00524; DWB; 1.
SQ SEQUENCE 423 AA; 47752 MW; 59960FC0E8C86A80 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTPPGY 6
Db 177 TTPPGY 182

Search completed: October 13, 2005, 15:09:29
Job time : 12.4286 secs

```

---